Gregory, Christian A., Ph.D. Three Essays on the Economics of Obesity. (2010)
Directed by Dr. Christopher J. Ruhm. 142pp.

The literature examining the relationship between obesity and wages has fairly consistently found that BMI has a negative impact on earnings for women, and less (if any) consequences for men. In the first study in my dissertation, co-authored with Christopher J. Ruhm, we relax the assumption–largely unquestioned in this research–that the conditional mean of wages is linear or piecewise linear in body mass index (BMI). Using data from the 1986 and 1999-2005 Panel Study of Income Dynamics, we estimate semi-parametric wage models that allow earnings to vary with BMI in a highly flexible manner. For women, the results show that earnings peak at levels far below the conventional threshold of obesity or even overweight. For men, our main estimates suggest a reasonably flat BMI-wage profile that peaks early in the "overweight" category. The findings for females (and the IV estimates for males) suggest that it is not obesity but rather some other factor – such as physical attractiveness – that may be producing the observed relationship between BMI and wages.

In the second essay of this dissertation, I examine the effect of obesity–and body mass more generally–on wages across the age distribution, using conventional parametric and more flexible semiparametric approaches. My parametric results suggest that the literature may overstate the effect of BMI and obesity on wages for women and almost certainly understates any negative association for men. For women, my results show that the negative effects of BMI and obesity are concentrated among women between 25 and 35 years old. While women in this age group experience an average 0.5 to 0.7 percent decrease in wages for each point increase in body mass (roughly 7.5 pounds), women over 40 will suffer a 0.25 percent decrease in wages for each extra point of body mass, and may not experience any wage penalty at all. Similarly, women who are 31-35 years old experience a 7.7 percent decrease in wages for being obese, while women over 40 experience only a 3.9 percent decrease. More flexible models largely confirm these results. For men, my parametric results indicate that, for those who are in their 20's or early 30's, BMI has no effect on

wages or is associated with a small increase; this is consonant with the rest of the literature. However, for men over 35, the effect of extra body mass is clearly negative: an extra BMI point brings with it a 0.3 percent decrease in wages for men 36-40 years old, and for men over 40, an extra BMI point is associated with a 0.5 percent decrease in wages. More flexible semiparametric models suggest that the negative association of BMI and wages may be as much as three times more than these estimates in some ranges of BMI. For instance, these models suggest that men 36-40 years old who have a BMI between 27 and 37 experience a 0.9 - 1.2 percent decrease in wages for each extra BMI point, as opposed to a .3 percent decrease predicted by the linear model.

Finally, in the third essay, I examine whether it has gotten easier to be obese over the last 25 years. Improvements in treatments for co-morbidities of obesity–high cholesterol, diabetes, sleep apnea and heart disease–over the last 25 years have made obesity less burdensome. I use data from the NHIS to examine whether such improvements have been borne out in obese persons' self-reports of health. My results suggest that between 1982 and 1996 obese women enjoyed significant gains in health relative to their normal weight counterparts. However, these gains do not appear to be due to improvement in treatment for co-morbidities of obesity; rather, income and especially education explain a large share of these health trends. For men, there seems to be little in the way of trends during these years of the survey. Results from the later years of the NHIS survey (1997-2006) suggest very little in the way of trends in self-reported health for obese men or women, but they suggest very large and significant improvements in health for obese women with coronary heart disease and obese male diabetics. All of these results should be interpreted with caution, as evidence of reporting anomalies in health appear to be present.

Three Essays on the Economics of Obesity

by

Christian A. Gregory

A Dissertation Submitted to
the Faculty of The Graduate School at
the University of North Carolina at Greensboro
in Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy

Greensboro
2010

Approved by

____

Committee Chair

To the memory of my mother and father.

Approval Page

This dissertation has been approved by the following committee of the Faculty of the Graduate School at the University of North Carolina at Greensboro.

Committee Chair ______________________________

Christopher J. Ruhm

Committee Members ______________________________

David C. Ribar

____

Peter M. Bearse

____

Christopher A. Swann

____

Date of Acceptance by Committee

____

Date of Final Oral Examination

ACKNOWLEDGEMENTS

I want to acknowledge my dissertation advisor, Christopher J. Ruhm, who always challenged me to learn more and be a better economist. The other members of my advisory committee–David C. Ribar, Peter M. Bearse, and Christopher A. Swann–have been indispensable to my progress. Each paid generous attention to my work: their suggestions have made my research better.

I would like to thank Michael Grossman, Naci Mojan, and the participants at the National Bureau of Economic Research Conference on the Economic Aspects of Obesity, held at Louisiana State University in November of 2008. Their responses and comments to early drafts of the first essay in this dissertation shaped and improved it.

Chapter I of this dissertation is due to be included in *Economic Aspects of Obesity*, ed. Michael Grossman and Naci Mojan, published by University of Chicago Press.

Two people outside the discipline of economics supported me and this work from start to finish. The first is my partner, Waits Raulerson, without whom this effort would have been impossible. The second is Sakyong Mipham J. Mukpo, who has been a source of constant inspiration and encouragement.

PREFACE

In the last 30 years, the proportion of the US population that is obese has more than doubled.[1] From a public health perspective, this increase is alarming because of the correlation of obesity with many debilitating and expensive co-morbidities: diabetes, some kinds of cancer, heart disease, stroke, hypertension, sleep apnea, and arthritis to name a few. From an economic perspective, this upward trend in obesity prevalence is less cause for alarm as occassion to examine the economic factors that contribute to obesity (and its increase) and the economic consequences of this condition.

The essays in this essay address these two aspects of the obesity. The last essay–"Has it Gotten Easier to be Obese?" addresses the first concern–the causes of the upward trend in obesity. It looks at whether any of this upward trend in obesity can be explained by the fact that treatments for the co-morbidities of obesity have made its health consequences less severe over the last 25 years. One of the disincentives to becoming obese is, of course, the expected health costs associated with it. But if treatments for, say, heart disease, are making it less likely that one will die of a heart attack even though one is still obese, what would be the incentive to change one's lifestyle, to lose weight and exercise? Although the results of this study are inconclusive, this is still an open and interesting question for further research: namely, is there a moral hazard effect from improved medical treatment?

The first two essays examine the effect of obesity–actually, body mass quite generally–in the labor market. In the first essay, co-authored with Christopher J. Ruhm, we look at this effect by using a statistical model heretofore not used in the literature. The intuition behind this approach is that we know relatively little about the mechanism for the effect of body mass index (BMI) on wages, and relying on the conventional BMI categories–underweight ($BMI < 18.5$), normal weight ($18.5 \leq BMI < 25$), overweight ($25 \leq BMI < 30$) and obese ($BMI \geq 30$)–in this context seems somewhat premature. Using a very flexible model, we show that, for women especially, the evidence to suggest that obesity or BMI is a stand-in for underlying health in wage determination is thin, at best. For men, our results are a

[1] Obesity is defined as having a body mass index–weight in kg/height in meters squared–greater than 30.

little more open ended, although we show that it seems likely that the effect of BMI on wages may greater than has been heretofore thought.

The second essay looks at the second concern mentioned above. It looks at the relationship between BMI and wages over the age distribution, with the idea that we should expect to see the penalty to obesity change with age if any of the mechanisms thought to drive this relationship are really at work. At a minimum, we should expect to see the effect of BMI increase as persons age if what is driving the effect of BMI on wages is health. My results suggest that this is less plausible for women than for men. Indeed, using methods similar to those I use in the first essay, I show that one neglected aspect of the discussion of the changing effect of BMI on wages is the BMI distribution itself, changes which seem to explain much of the change in the average effect of BMI on wages as persons age. This confirms the impression that health is not driving the relationship between wages and BMI for women, and, although this is more plausible for men, it is still an interpretation that still needs more empirical work to support.

These essays open up many further avenues for further research, including into the relationship between health conditions, BMI and wages, and into measurement models for self-reported health. These are important for economists, especially if we want to understand and make sound policy concerning health outcomes, particularly those related to body mass.

TABLE OF CONTENTS

LIST OF TABLES

Page

LIST OF FIGURES

CHAPTER I

WHERE DOES THE WAGE PENALTY BITE?

(Co-authored with Christopher J. Ruhm)

Abstract

The literature examining the relationship between body mass index (BMI) and wages has fairly consistently found that BMI has a negative impact on earnings for women, and less (if any) consequences for men. In this paper, we relax the assumption–largely unquestioned in this research–that the conditional mean of wages is linear or piecewise linear in body mass index (BMI). Using data from the 1986 and 1999-2005 Panel Study of Income Dynamics, we estimate semi-parametric wage models that allow earnings to vary with BMI in a highly flexible manner. For women, the results show that earnings peak at levels far below the clinical threshold of "obesity" or even "overweight". For men, our main estimates suggest a reasonably flat BMI-wage profile that peaks early in the "overweight" category. However, the results of instrumental variables (IV) models or specifications focusing on log lags of BMI are more similar to those for women. The findings for females (and the IV estimates for males) suggest that it is not obesity but rather some other factor – such as physical attractiveness – that may be producing the observed relationship between BMI and wages. We also provide non-parametric estimates of the association between BMI and health expenditures, using data from the Medical Expenditure Panel Survey. These cast further doubt on the hypothesis that the wage penalties associated with increasing BMI occur because the latter serve as an index for underlying medical costs.

Introduction

How does BMI affect wages? At first blush, the answer seems obvious. Over the last 15 years, a large literature has established the negative correlation between obesity–the condition of having a body mass index (BMI) greater than 30–and wages, at least for women. On average, obese women make two to eight percent less than their normal weight counterparts. Obese men don't make any less than men of normal weight, and heavy black men may earn slightly more.[1]

The question we ask is not about obesity, however, at least not obesity alone. We are interested in the more general relationship between BMI and wages. In particular, we examine two assumptions that characterize previous research. The first is that the BMI range above 30 is "where the action is." Although there are good reasons to focus on obese persons, the rest of BMI distribution has been treated as an afterthought in most of this literature. The second is that the conditional expectation of wages is linear in BMI, or characterized by some other relatively simple parametric relationship (such as a quadratic). While specifications based on these assumptions are valuable because they are tractable and easily interpretable, there are good reasons to assume they are not true *ex ante*. In the simplest case, if BMI really does reflect something meaningful about health, it could be that wages are negatively associated with both overweight and underweight. Linear models capture only the average effect–which, in this example, might well be zero–and therefore miss important ways that BMI affects earnings.

Only recently have economists begun examine the shape of the conditional wage function. Wada and Tekin (2007) is the first study we are aware of that allowed a measure of body weight to enter into a wage regression as a quadratic. Even more recent has been the adoption of semi-parametric methods. Shimokawa (2008) used data from China to estimate semi-parametric models and finds that wages are lower for men and women in the tails of

[1] Throughout, we use the conventional definitions of "underweight," "healthy" (or normal) weight, "overweight" and "obese" for persons in the BMI ranges of: <18.5, 18.5-<25.0, 25.0-<30.0 and ≥ 30.0 (National Heart, Lung and Blood Institute, 1998).

the BMI distribution. Kline and Tobias (2008), using data from the 1970 British Cohort study, found that marginal increases in BMI are most harmful for men who are overweight or obese and for women in the "healthy" weight range.

In addition to examining the shape of the conditional wage function, we address potential biases resulting from endogeneity of BMI and possible reverse causation, whereby wages determine body weight. We deal with endogeneity using an instrumental variables (IV) approach, where the respondent's BMI is instrumented with sibling BMI. To address the potential problem of reverse causality, we follow previous research in using lagged body weight to rule out the effect of current wages on weight. However, our analysis employs longer lags (at least 13 years) and BMI from relatively early in the typical worklife. Both general approaches have been used before but ours is the first application on data for U.S. subjects using semi-parametric (SPM) methods.

We also examine potential mechanisms by which BMI affects wages and, in particular, are interested in understanding gender differences in these effects. Researchers have pursued several possibilities in this regard. One is that body weight affects health expenditures for women in a way that it does not for men, and that overweight and obese women pay for these expected expenditures in the form of lower wages (Bhattacharya and Bundorf, 2005). Another is that health differences due to obesity have disparate effects on marginal productivity (Baum and Ford, 2004). Still another is that women working in professions requiring public interaction are more penalized for obesity than corresponding men (Baum and Ford, 2004; Pagan and Davila, 1997; Han et al., 2009b). Or, finally, employers might discriminate against overweight or obese women but not men. Although direct evidence is only provided on the first of these possibilities, we interpret our findings in context of the growing literature examining how "beauty" is related to earnings.

Our analysis produces three main results. First, women's wages peak at thresholds far below the obesity cutoff, usually at a BMI of 23 or lower. This finding is robust to specifications correcting for endogeneity or reverse causation and suggests that BMI does not serve as an index of underlying health or medical costs in a wage-setting context. We test and confirm this intuition through a non-parametric analysis of relationship between BMI and medical expenditures. An alternative, which we believe to be more consistent with our findings, is that BMI is a proxy for physical “attractiveness” (or “beauty”), which is known to affect earnings.

Second, the estimates for men are more dependent on the choice of preferred models. Our primary specifications suggest that the conditional wage function is increasing in BMI through the beginning of the range of “overweight” and remains constant or declines modestly thereafter. Conversely, models using long-lags of BMI or instrumental variables indicate that male wages peak at very low BMI levels, suggesting that, as for women, the observed patterns are more likely to indicate physical attractiveness than underlying health status or medical costs.

Third, there are often substantial differences for blacks and whites, with the main specifications suggesting that the conditional wage function peaks at a considerably higher BMI for minorities and declines more slowly thereafter. Such findings might be consistent with a role for attractiveness, if there are racial differences in perceptions of ideal body weight. However, the IV estimates reveal smaller racial disparities, so that these interpretations require caution.

Data

We use data on 25-55 year olds from the 1986, 1999, 2001, 2003, and 2005 waves of the Panel Study of Income Dynamics (PSID), a longitudinal survey that began in 1968 with 4,802 families.[2] An additional 581 immigrant families were added in 1997 and 1999, and new families were created from the existing ones due to the formation of new households (e.g. due to divorce or to grown children leaving home).[3] As of 2005, the PSID contained 8,041 families.

[2] The original sample includes a nationally representative group of 2930 families, with the complement from a low-income sample.

[3] An earlier attempt to include Latino immigrants dates to 1990, at which time 2,043 immigrant families from the three most prevalent Latino groups in the United States were included. This sample was dropped after 1995.

Previous related studies involving U.S. subjects have used data from the National Longitudinal Survey of Youth 1979 (NLSY). We chose instead to utilize the PSID, primarily because it has characteristics of both longitudinal and cross-sectional data. Since the NLSY provides information for a single fairly narrow birth cohort covering a somewhat limited age range, previous analyses using it have been largely restricted to relatively young workers. By contrast, the PSID is a self-replenishing panel that began in 1968 and so is more suitable to addressing differences in the effects across age groups. As we argue later, such differences point to possible mechanisms by which BMI affects earnings. That said, we show that our results are not driven by use of the PSID sample: similar patterns are obtained using comparable age ranges in the PSID and NLSY.

The PSID gathers information through an interview with one primary adult–usually the male head of household, referred to as the "head". On occasion, the spouse or cohabiting partner, the wife/"wife," as she is called, is the family respondent. In the waves used for this study, the PSID collects data on height and weight of the head and wife/"wife" only. The survey respondent gives height and weight information about themselves as well as their spouse or cohabiting partner. In an effort to minimize reporting error, we include only observations for which the head or wife reports his/her own height and weight.

Self-reported height and weight contain errors. We adjust for these using the regression correction suggested by Lee and Sepanski (1995) and commonly employed in the literature (Cawley, 2004; Chou et al., 2004; Lakdawalla and Philipson, 2007). Specifically, using data from the National Health and Nutrition Examination Survey (NHANES) III (1986-94), NHANES 1999, NHANES 2001, and NHANES 2003, we regress measured height (weight) on self-reported height (weight), its square and its cube. The results, for models stratified by gender and race, are used to predict actual BMI (in the PSID) as a function of self-reported BMI.[4]

Hourly wages are constructed by dividing total earnings for the calendar year previous

[4]We use multiple waves of NHANES so that we can restrict the age range of the prediction samples to those relevant to our earnings study: namely, persons 25-55 years old.

to the interview by total hours worked in that year.[5] For all but a handful of persons, total earnings and hours refer to the main job: very few people report second jobs or overtime earnings. The PSID imputes wages for people who report earnings but not hours or vice versa. We retain these observations (less than 2 percent of our sample) although our results are not sensitive to doing so. Our sample includes 25 to 55 year olds who worked at least 20 hours per week in their main job. These restrictions limit the sample to prime-age workers. We normalize wages to 2005 dollars using the CPI, drop observations reporting wages less than half of the federal minimum, and trim the top $\frac{1}{2}$% percent of wage observations.[6] Our final analysis sample contains 7,251 person-years for women and 5,775 person-years for men. Among women, we observe 1,433, 1,095, 516, and 520 persons in 1, 2, 3, and 4 years, respectively. Among men, we observe 1,007, 666, 424, and 541 persons in 1, 2, 3, and 4 years, respectively.

Methods

The estimates were obtained using a semiparametric (SPM) local linear regression framework that can be usefully distinguished from both OLS and a univariate kernel regression model. As is well known, ordinary least squares assumes that the conditional mean of the dependent variable is a linear function of the independent variables. This makes it easy to make predictions and to gauge statistical significance of the coefficients. However, the assumption of linearity is restrictive in ways that can only partially be overcome through standard transformations, such as including higher order polynomials of the explanatory variables of key interest. Kernel regression drops the linearity assumption and instead models the expectation of the dependent variable as a weighted mean at every point in the

[5] Validity of the PSID income and hours data has been repeatedly evaluated. In two of the most cited evaluations (Bound et al., 1994; Duncan and Hill, 1985) earnings were found to be relatively free from reporting error, but work hours were subject to significant mistakes. This induces errors into hourly earnings unlikely to abide by textbook assumptions about correlations between these variables and key regressors. However, there is no reason to believe that work hours in the PSID are subject to more reporting mistakes than similar measures in other data sets such as the Current Population Survey or NLSY (Bound et al., 2001; Hill, 1992).

[6] This procedure drops women with a wage above $75.14 and menn with a wage higher than $152.57.

distribution of the independent variable. While this model can produce accurate univariate estimates with relatively small samples, in multivariate settings, it is not possible to maintain a meaningful level of accuracy without the sample size increasing exponentially. In these contexts, we use the specification

$$Y_i = z_i * \beta + f(BMI_i) + \varepsilon_i, \tag{1.1}$$

where Y_i is hourly wages of individual i, z_i is a vector individual characteristics and year effects, and $f(BMI)$ is the non-parametric function transforming BMI into wages, which we refer to as the "conditional wage function."[7] The resulting models are semi-parametric because they assume that the covariates included in z are linearly related to wages, whereas flexibilty is maintained in transforming BMI into earnings.

Our estimates use the stepwise "double residual" method outlined in Robinson (1988). In the first step, we estimate $\hat{Y}_i$ and $\hat{z}_i$, as predicted values from a non-parametric regression of each of the independent and dependent variables on BMI. From these we derive $e\hat{p}s_i^Y = Y_i - \hat{Y}_i$ and $e\hat{p}s_i^z = z_i - \hat{z}_i$, representing the portions of the dependent and explanatory variables that are unrelated to BMI. In the second step, we regress $e\hat{p}s^Y$ on $e\hat{p}s^z$ to get $\hat{\beta}_{eps}$. Finally, we estimate the conditional wage function, $\hat{f}(BMI_i)$, by non-parametrically regressing the wage residual $Y_i - z_i * \hat{\beta}_{eps}$ on BMI_i, using the techniques detailed in the Appendix.[8] The intuition behind this procedure is to purge the dependent variable of the portion of the supplemental variables that are unrelated to BMI and then provide a local linear regression estimate showing the relationship of this residual to BMI itself. We estimate confidence intervals using the "wild" bootstrap algorithm outlined by Yatchew (1998, p. 688) and Yatchew (2003, pp. 160ff).[9]

[7] We use levels instead of logarithms of wages to make our estimates easily interpretable in the figures and tables. Using log wages as the dependent variable yields quantitatively and qualitatively similar results.

[8] We also estimated $f(BMI)$ using the first differencing procedure outlined by Yatchew (2003), and obtained essentially the same results. However, we maintained the double residual method for our point estimates and confidence intervals to preserve efficiency.

[9] This algorithm is often applied when heteroskedasticity is a concern. To form 95-percent confidence intervals, we resample 1200 times from the residuals to form bootstrap data sets and perform the local linear regression procedure outlined in the Appendix at between 200 and 300 points in the BMI distribution.

For our instrumental variables estimates, we use the same stepwise procedure, but add to the first-stage the residuals of a linear regression of BMI on the instruments. Just as with the other explanatory variables, we form a non-parametric prediction of the residual conditional on BMI ($i\hat{veps}$) and a residual ($e\hat{p}s^{iveps}$). We include that residual in the second stage residual regression and form our estimate of $\hat{f}(BMI)$ as above. This procedure removes the variation in BMI *not* explained by the instruments from the second stage regression, so that what identifies $\hat{f}(BMI)$ is what the instruments do explain (Shimokawa, 2008; Yatchew, 2003).

We employ two strategies to address the problems that hamper estimation of the causal effect of BMI on earnings. First, to deal with the issue of reverse causality, we estimate models in which the independent variable of interest is lagged BMI (see Sargent and Blanchflower, 1994; Averett and Korenmann, 1996; Baum and Ford, 2004; Cawley, 2004). The general argument subtending this strategy is that current wages might influence current BMI but cannot affect BMI in previous years. However, a statistical association may exist if body weight or wages are correlated across time. We address this difficulty in two ways. First, where previous related studies have used BMI lags of up to 7 years, we analyze wages in 1999-2005 as a function of BMI in 1986, or 13-19 years earlier. Second, we limit this portion of the analysis to individuals less than 26 years old in 1986, under the assumption that wages early in the person's work career are unlikely to determine BMI during middle-adulthood.

To account for the potential endogeneity between BMI and wages, we follow an instrumental variables strategy similar to that developed by Behrman and Rosenzweig (2001), and more recently used by Cawley (2004), where sibling BMI is the instrument.[10] The validity of this strategy rests on the suppositions that sibling BMI is correlated with own BMI and that it is uncorrelated with one's own earnings, except through BMI. The first assumption

[10]Kline and Tobias (2008) have similarly used parent BMI as an instrument; Shimokawa (2008) has used sibling BMI and lagged child weight as instruments. An alternative is to estimate fixed-effects (FE) models (Baum and Ford, 2004), which automatically account for all time-invariant sources of heterogeneity. However, FE methods may be problematic for this application because they assume that weight changes translate instantly (or very rapidly) into wage changes, whereas current earnings are likely to be affected by both contemporaneous and past body weight.

is uncontroversial and can be tested. The second is more problematic. In particular, sibling BMI could be independently related to wages if siblings share traits affecting both weight and wage outcomes due to environmental influences or genetics.

Until recently, much of the literature suggested that the environmental influences on body weight tend to be non-shared between siblings, and that their importance diminishes in adolescence (Maes et al., 1997). However, recent developments suggest that environment may be more important than once thought.[11] Similarly, the emerging literature linking genetics to human behavior suggests caution. For example, certain polymorphisms of the D4 Dopamine receptor gene are correlated with attention-deficit hyperactivity disorder (Sunohara et al., 2000; El-Faddagh et al., 2004).[12] It is well known that the regulation of dopamine affects experiences of satiation and therefore eating behavior.[13] Research has also found that both childhood inattention and adult obesity are correlated with the Dopamine D4 receptor gene in women with Seasonal Affective Disorder (Levitan et al., 2004). These studies raise the possibility that child behaviors affecting learning and subsequently wages may be correlated with genetic factors also influencing body weight.[14] Therefore, care is needed in interpreting the results of IV models (like those below) identified by genetic variation in BMI.

[11] Most studies attribute the effect of genetics to the difference in the covariance between monozygotic (MZ) and dizygotic (DZ) twins' body weight, since DZ twins share only half their genetic material with the other twin. But in addition to having different genes, DZ twins may also have different dominant and recessive copies of shared genes. This "non-additive" genotype variation might explain a significant amount of variation in traits such as body weight. One recent study (Segal and Allison, 2002) identifying this variation through the use of "virtual twins"–same-aged siblings that don't share any genetic material–found that a 5 to 45 percent of the variation in BMI could be due to environmental influences.

[12] Swanson et al. (2000) found no correlation between the presence of the genetic trait and neuropsychological abnormalities sometimes associated with ADHD; however, they did find a correlation between the genetic marker and extreme behavior.

[13] However, at least one study failed to find a direct direct link between obesity and the D4 dopamine receptor gene (Poston et al., 1998).

[14] Holtkamp et al. (2004) found that children with ADHD were also more likely to be obese, suggesting the plausibility of a genetic connection.

Full Sample Results

We next summarize our semi-parametric estimates of the relationship between BMI and wages. Throughout, we stratify by sex, since BMI could have quite different effects for men and women.[15] All models control for age, marital status, number of children, presence of a child less than two years old in the household, level of schooling, job tenure (in months), the survey year, and region of residence.[16] Race/ethnicity are also held constant in the full sample estimates (but not when stratifying by race). Unless otherwise noted, the y-axis of the figures indicates the expected wage, calculated by adding $\hat{f}(BMI)$ to the group-specific average predicted wage; results are displayed for BMI ranging from 20 to 40.[17]

Main Specifications

Figure 1.1 shows full sample estimates. The conditional wage function of women is characterized by a peak at a BMI of 22.8. Weight gains at lower BMI are associated with higher earnings, although the confidence intervals are sufficiently large that we can not generally reject the null hypothesis of no effect. By contrast, predicted wages decline rapidly at higher BMI levels, and monotonically, expect for a statistically insignificant upwards tick just below the obesity threshold.

These findings suggest that female wages begin to fall well before conventional cutoffs for "obesity" or "overweight", and even well within the "healthy" weight range. Thus, there is little evidence of an obesity penalty per se. Instead, the data suggest that women whose weight rises above a relatively low threshold experience reduced earnings. Of course,

[15] All estimates are unweighted, in part because the PSID assigns a zero weight to anyone entering the sample through co-habitation or marriage. To ensure that our results are not driven by this choice, we estimated models using only the nationally representative sample or limiting the analysis to observations with positive weights and using these weights in the second-stage regression (of $e\hat{p}s^Y$ on $e\hat{p}s^z$). In both cases, the results are essentially the same as those shown.

[16] We excluded occupation from our primary estimates, since this is one mechanism through which BMI could affect earnings. Specifications adding controls for broad occupational categories resulted in similar estimates for women and flatter BMI-earnings profiles for men.

[17] This range covers approximately the 5th through 95th percentiles of women and the 1st through 98th percentiles of men. We exclude from the analysis persons with BMI greater than 45, as these observations exert disproportionate influence on the semi-parametric estmates. This trimming drops 34 men and 125 women.

BMI does not perfectly measure obesity and some women in the "normal" BMI range may actually be clinically obese.[18] However, even if there are classification errors, the very low BMI at which the wage function peaks makes it much more probable that we are observing the effects of appearance or beauty, rather than obesity or poor health. A growing literature suggests that attractive individuals earn more than their counterparts (Biddle and Hamermesh, 1994; Hamermesh and Biddle, 1998; Harper, 2000; French, 2002), although the mechanisms for this are not fully understood. A possible explanation for our results is that females are considered most attractive at low levels of BMI. Consistent with this, Maynard et al. (2006) provide evidence that the desired BMI of adult women is between 22 and 23, or almost exactly where the conditional wage function peaks.

The patterns for men differ substantially. Predicted wages are maximized at a BMI of 26.7 – in the "overweight" range – with lower and higher bodyweight associated with substantial but imprecisely estimated decreases. Yet these results also provide little evidence of a sizeable "obesity penalty", except perhaps at extremely high BMI. Instead, they raise the possibility of wage reductions from being too light. For instance, the predicted hourly wage of a man with a BMI of 35 is just \$0.81 per hour below that of his peer with a BMI of 27, while a BMI of 20 is associated with hourly earnings that are \$3.19 less. Such results are consistent with the possibility, supported by previous evidence (DiGioachino et al., 2001; Maynard et al., 2006) that males are held to a different appearance standard than females, with "thin" women viewed as attractive while corresponding men are considered "scrawny." However, as discussed below, we obtain considerably different estimates for men (but not women) when using instrumental variables techniques, so these results should be interpreted with some caution.

Are Semi-Parametric Estimates Worth the Effort?

Are the benefits from using the semi-parametric models are worth the added complexity (and computational time) need to estimate them? Our answer is a qualified "yes." To illustrate

[18] Burkhauser and Cawley (2008a) provide evidence that BMI is more likely to understate than to overstate obesity prevalence.

the potential gains from these estimates, Figure 1.2 plots the results from modeling wages as linear or quadratic in BMI, alongside the SPM estimates that are novel to this analysis. The conditional wage function of women is monotonically decreasing in BMI for the linear and quadratic specifications, which provide essentially identical estimates. While generally reasonable, the parametric models miss the increase in the wages occuring below a BMI of 23 (although the differences are small and often not significant), and understate the drop in earnings predicted immediately thereafter. At the very least, the SPM estimates suggest that the conditional wage function is flat until a BMI of 23, and decreasing nearly monotonically thereafter.

For men, the gains to more flexible models are larger. In Figure 1.2, it is clear that the linear specification fares the worst. The quadratic model does better in approximating the conditional wage function, and is sensible if we think that health effects or costs of obesity drive the BMI-wage relationship and begin to bind the wage function at *some* point in the BMI distribution. However, even the quadratic model is restrictive – overestimating wages at low BMI and in the "overweight" range, and indicating that the conditional wage function is maximized at a considerably higher BMI than the semi-parametric model. These differences are non-trivial since the quadratic specification suggests an "obesity penalty," while the more flexible estimates indicate that wages begin to decline much earlier, indicating that other factors may be at work.

Potentially useful, and computationally cheaper, alternatives to our SPM procedure might involve estimating models with higher order polynomials in BMI or linear splines.[19] Indeed, we would recommend these as time-efficient and relatively simple procedures for much future research. However, the preferred parametric specification may not be obvious *a priori*. The semi-parametric procedures employed here may help to guide that choice and provide a more complete understanding of the conditional earnings function.

[19] For example, `Stata` has a pre-programmed routine (the `lpoly` command) that will estimate local polynomial fits with usable, although not asymptotically correct, confidence intervals.

PSID vs. NLSY

Previous related U.S. research has generally used data from the NLSY, rather than the PSID. Although we view the PSID to be preferable in several respects, most importantly because it is not limited to a single cohort or narrow age range, we checked whether the results were sensitive to its use. To do so, we obtained NSLY data for 1998 through 2004 (approximating the years of our main PSID analysis), during which time NLSY respondents were 33 to 47 years old. We constructed a sample of correspondingly aged individuals from the PSID and performed two analyses. First, we estimated simple OLS models for the two data sets.[20] For women, the estimates turned out to be quite similar. For instance, the coefficient (standard error) on BMI was -0.122 (0.017) in the PSID and -.168 (.024) in the NLSY.[21] For men, the results were somewhat different: using the PSID, we obtained a coefficient (standard error) of 0.017 (0.044), while the estimates were -.192 (.043) for the NLSY. The PSID findings are consistent with those shown in Figure 1.2. Although the NLSY estimates for males run counter to some prior research (which does not uncover an obesity effect on wages), this is likely due to the young age range of the men previously examined. Gregory (2010) and Han et al. (2009b) have recently shown that the negative correlation between BMI and wages strengthens as men age, consistent with our results.

Second, we ran semi-parametric models for the PSID and NLSY subsamples. These estimates, summarized in Figure 1.3, reveal generally similar patterns.[22] However, there is evidence of greater non-linearities for women in the PSID than the NLSY, while the male wage function reaches a maximum at a lower BMI in the NLSY. Overall, it seems likely that we would find even less evidence of a pure "obesity effect" in the NLSY, since the conditional wage function is maximized at a lower BMI. However, since the female wage function is approcximately linear in the NLSY, there might be less gain from the flexible

[20] The NLSY data include only persons in the representative sample and we use similar sample restrictions as in the PSID. The regressions are not weighted. Since we cannot easily identify pregnant women in the PSID, we run specifications for the NLSY data with pregnant women included. Separate NLSY models that exclude pregnant women yield similar results.

[21] Our results are also similar to those obtained by Cawley (2004), when we estimate models using the log (rather than level) of earnings, as he did.

[22] The smoothing estimates were normed to address some differences in scaling between the two data sets.

SPM estimates.

Reverse Causation

The preceeding findings could be biased due to reverse causation, where higher wages lead to lower BMI. For example, this could occur because high-earners can more easily afford expensive foods, such as fruits and produce, that are healthy and low in calories. Alternatively, they may have greater flexibility in their jobs to find time to exercise and could more often join health clubs. We examine this issue in Figure 1.4, which shows how lagged BMI is related to wages. Specifically, we measure BMI in 1986 and wages during 1999-2005. To reduce the possibility that lagged BMI itself is strongly influenced by (prior) earnings, we restrict this analysis to persons less than 26 years old in 1986, and so at the beginning of their worklives. Since BMI typically rises with age, the distribution of lagged BMI is to the left of the contemporaneous distribution. Therefore, Figure 1.4 displays BMI (in 1986) over the range 18 to 37, rather than 20 to 40.[23]

The results for long-lags of BMI and are fairly similar to those using contemporaneous weight (and the full sample), once we account for the lower average BMI of young adults, and they again provide scant evidence of an "obesity penalty." Specifically, the female wage function peaks at a very low BMI level (below 18) that is actually in the "underweight" category, although the earnings penalties thereafter are not always monotonic or statistically significant. For men, lagged BMI is essentially unrelated to contemporaneous wages, but with the peak predicted at a very low (18.6) BMI. These patterns are similar to those of women and suggest that being "thinner" is (almost always) better for males as well as females. We return to this result when examining our instrumental variables estimates.

Instrumental Variables

BMI could be correlated with unobserved factors also affecting wages. For example, persons earning high wages because they are motivated at work might similarly be motivated to

[23] This corresponds to approximately the 5th to 96th percentile of the female BMI distribution in 1986.

exercise and consume healthy diets. The same might be true for individuals with low discount rates. In both of these cases, BMI will be correlated with the error term in our wage specification. We address this possibility by estimating instrumental variables estimates, using sibling BMI as the instrument.[24] These results are shown in Figure 1.5.

For women, the IV estimates are similar to those obtained in the main models. Specifically, the conditional wage function is maximized at an even lower level of BMI (21.4), with a rapid decline in earnings predicted from the middle of the "healthy" weight range to just beyond the threshold for "overweight". However, the wage function is flat after a BMI of 26, further suggesting that we are not observing the effects of obesity.

IV estimation makes a much larger difference for men. Where the main specifications indicated that the wage function increased into the "overweight" range, and then declined relatively slowly, the IV models suggest essentially no effect through a BMI of 25 or so but with wages predicted to fall rapidly thereafter. Such results could indicate a role of poor health or medical costs but only if the effects begin to bind at the beginning of the "overweight" category. This seems unlikely, since most available research (Quesenberry et al., 1998; Andreyeva et al., 2004; Arterburn et al., 2005), suggests that health costs are similar for "healthy" weight and "overweight" individuals but substantially higher for obese and, especially, severely obese persons.

Race

The wage functions of white and black females differ markedly (see Figure 1.6). As in the full sample, the earnings of white women are predicted to peak well below the "overweight" threshold (at a BMI of 22.5), to decline markedly immediately thereafter, but then to be relatively flat beyond the middle of the "overweight" catgory. By contrast, the pattern for black women is consistent with a true "obesity penalty", since the maximum predicted wage occurs at a BMI of 26.1 and nearly all of the economically or statistically significant

[24] In a standard linear model, first-stage F-statistics on the instruments are 29.5 for women and 16.2 for men, well in excess of the level of 10 recommended by Staiger and Stock (1997) to avoid problems with weak instruments.

reduction takes place at or beyond the obesity threshold. However, these results probably do not indicate that the obesity effect is due to higher medical costs or health problems. Were this the case, we would expect the wages of severely obese individuals to be substantially below those of their mildly obese counterparts (since severe obesity has by far the most deleterious health consequences). Instead, there is no evidence that the wage function continues to decline beyond a BMI of 35.

The results for men are even more interesting. The wage function of white males reaches a maximum at a BMI of 26 but remains relatively flat subsequently, with even severely obese men predicted to earn only modestly less. Conversely, the expected earnings of black males rise well past the obesity threshold (to a BMI of 32.1) and then remain flat or decline modestly.

These findings suggest substantial race differences in the BMI-wage profile, with greater and more binding weight penalties for whites than blacks that, except for black men, begin well before the obesity threshold.[25] Assuming that the relationship between BMI and health or medical costs is similar for blacks and whites, the racial disparities make it unlikely that the results in Figures 1.6 and 1.7 reflect underlying effects of BMI on health conditions or medical costs. Instead, we think it more probable that these reflect appearance effects, combined with different standards of "desired weight" being applied to blacks and whites (and males and females).[26]

Simulations

Table 1 displays semi-parametric estimates of the difference in predicted wages at specified BMI levels, relative to a reference group of females with a BMI of 23 or males with a BMI of 27.[27] The results are presented for subsamples, stratified by race and sex, for both our main SPM specifications (using actual BMI) as well as from semi-parametric instrumental

[25]Instrumental variables suggest that this may also be the case for black males, as discussed below.

[26]For example, college students report higher "desired BMI" for African-American than white females and for females than males (DiGioachino et al., 2001).

[27]The reference category is chosen to approximate the BMI level maximizing the conditional wage function in the main full sample specifications.

variables (SPM-IV) models. Standard errors are estimated from bootstrap replications, with p-values assigned using the percentile method. Coefficient estimates for the supplementary regressors are contained in Appendix Tables A-1 and A-2.

Table 1 highlights several points made previously, as well as some new ones. First, the wage function for females begins to decline at a relatively low bodyweight. Compared to women with a BMI of 23, BMIs of 25, 30 and 35 predict statistically significant penalties of $0.96, $1.51 and $2.62 per hour. This pattern is driven by white females, where the conditional wage function indicates even larger (although less precisely estimated) gaps of $1.02, $1.93 and $3.51 per hour. The IV models reveal a similar pattern for white women, although with somewhat weaker predicted wage declines and standard errors that "blow up" at BMIs above 35. Conversely, the findings for black females are more dependent on the choice of estimation techniques. Using actual BMI, predicted earnings reach a maximum at a BMI slightly above 26 and then decline relatively slowly. However, the IV estimates suggest a flatter conditional wage function prior to the peak, which occurs earlier (at a BMI of 21.6), and with a more rapid decline thereafter. Thus, the IV estimates for black females look relatively similar to the patterns seen for white women.

For men, the primary SPM estimates suggest that only a small wage penalty is associated with high BMI, except perhaps for severe obesity. Thus, a BMI of 30 or 35 predicts hourly wages that are a statistically insignificant $0.21 and and $0.81 lower than expected at a BMI of 27, with larger gaps for white males but positive predicted effects for blacks. On the other hand, hourly earnings are anticipated to be two to four dollars lower at a BMI of 20 than for the reference group.

The IV results for males are quite different: the wage function is monotonomically downward sloping beginning at low levels of BMI, with very large penalties associated with excess weight. Thus, men at the obesity threshold (BMI=30) are anticipated to earn over four dollars per hour less than their counterparts with a BMI of 20; those with a BMI of 35 are predicted to receive about eight dollars less. These differences are of similar size for white and black men, with the most important disparity being that the conditional wage function declines substantially between a BMI of 20 and 25 for blacks, and then flattens temporarily, whereas the pattern is reversed for whites.

BMI and Medical Expenses

Obese individuals might suffer a wage penalty because they have high medical costs that are partially paid by employers, through the health insurance system. Bhattacharya and Bundorf (2005) offer a version of this argument, providing evidence from the Medical Expenditure Panel Survey (MEPS) that the wage effects of obesity, for women, are borne entirely by those with employer-provided health insurance and, further, that the expected health costs of obesity are significantly higher for women than men.[28] Based on this, they claim that the effect of obesity on female wages is due to employers who offer insurance trading off wages against expected health expenditures, rather than because of any "beauty premium" or "appearance penalty."

We are doubtful of such a mechanism for the simple reason that the conditional wage function for women turns downwards so early – at a BMI of under 23 – far below either the obesity threshold or the level at which health costs might be expected to increase. Nevertheless, we directly test the possibility that health expenditures explain our results in two ways. First, we use MEPS data to produce a univariate non-parametric estimate of the log of total health expenditures (in 2005 dollars) as a function of BMI.[29] If our previous results are explained by employers using body weight to risk-rate employees, we would expect the pattern of medical expenditures to approximately track that for earnings. In particular, the medical costs of women should begin to rise at low BMI, starting at around 23. The health expenditures for men should either not increase much prior to the obesity threshold (if we believe the results based on actual BMI), or show a similar pattern as for women, although starting to rise slightly later (if we place greater trust in the IV estimates).

[28] However, somewhat contradictory findings are obtained by Baum and Ford (2004).

[29] We used data from the MEPS 1999, 2001, 2003 and 2005 samples and trimmed the top 1% of BMI observations. Using levels, rather than logs, of expenditures gives similar results.

Figure 1.8 displays the non-parametric relationship between BMI and log medical costs.[30] For women, predicted health expenditures change little prior to the the obesity threshold but increase rapidly thereafter. This pattern is quite plausible but almost certainly indicates that medical costs do *not* explain the observed conditional wage function, since earnings begin to fall much earlier – in a region where body weight is essentially unrelated to health costs. By contrast, we observe a monotonically increasing BMI-medical cost gradient for men, which has some potential for explaining the wage function obtained from the IV estimates (but less so when using actual BMI).

Second, we examine how the conditional wage function varies with BMI for subgroups stratified by age and gender. The medical costs of obesity are likely to increase with age (Finkelstein et al., 2007). If such expenditures are the source of the fall-off in wages, we should therefore expect, *ceteris paribus*, a steeper BMI-wage gradient for older than younger persons. Instead, figure 1.9 shows that the conditional wage function declines from its peak much more rapidly for 35-44 than for 45-55 year old women. Similarly, wages are essentially unrelated to BMI for the oldest (45-55 year old) males, whereas the data suggest earnings penalties at high (and low) BMI for younger men (see figure 1.10). Finally, note that female wages are predicted to reach a maximum at a BMI of around 22 or 23 for all three age groups, well below the "obesity" or "overweight" thresholds. This seems inconsistent with the possibility that health expenditures are the primary determinant of the relationship between earnings and BMI.[31]

[30] Our analysis does not account for two important characteristics of the expenditure data. First, there are a lot of zeros: in our sample, accounting for roughly 12% (29%) of women (men). Second, the distribution is extremely skewed. A more appropriate specification, in a semi-parametric context, would be a partial general linear model using a gamma distribution and a log link (e.g. see Müller (2001)). However, such models are computationally expensive, even for parsimonious specifications, and we leave it to future research to explore the benefits of using them to examine the relationship between health expenditures and BMI.

[31] It is less clear what age-pattern is expected if "beauty" play a key role. If BMI becomes less closely tied to perceptions of beauty at higher ages, or if appearance itself becomes a less important determinant of wages, we would expect a steeper wage function for younger than older women. Conversely, appearance at young ages could have long-lasting consequences by directly influencing future productivity through, for example, its effects on self-esteem (Mobius and Rosenblat, 2006; Mocan and Tekin, 2006), or if initial labor market opportunities establish a path for future outcomes.

Discussion

The preceding analysis used semi-parametric regression methods to examine how body weight is related to wages. Compared to previous research, these specifications allow great flexibility on the role of BMI, while imposing standard parametric restrictions on the other included controls.

A particularly striking finding is that increased BMI is associated with wage reductions for white females, beginning at low levels of weight – considerably below conventional thresholds for "obesity" or "overweight." These results are robust to accounting for reverse causation or endogeneity and indicate that the conditional wage function is probably not being driven by the health effects of BMI or by obesity per se. Instead, they suggest that, over most of the BMI distribution, being "thinner is better" for white women, possibly due to social perceptions of beauty or desired appearance. The evidence for black females is more ambiguous. Our main specifications, conditioning on actual BMI, indicate that the earnings profile is flat prior to a BMI of around 26 but then begins to decline fairly rapidly. This might reflect a different appearance standard for nonwhites but also raises the possibility of an "obesity penalty" for this group.[32] However, instrumental variables estimates show a pattern more similar to that for white females, with earnings predicted to be maximized at a low BMI (21.8) and to decline rapidly thereafter.

The results for men are even more dependent on the estimation technique. In our main specifications, earnings increase through a BMI of around 27 and then fall modestly. Conversely, the IV findings look similar to those for women, in predicting that wages decrease with BMI throughout virtually the entire range of the latter. Controlling for reverse causation (by including long-lags of BMI) also yields a conditional wage function that is maximized at a low BMI level and is fairly flat thereafter. The findings for black males differ from corresponding whites in that the main (non-instrumented) specifications show an increase in the conditional wage function until well into the "obesity" range but with a

[32]For example, Stearns (1997) and Averett and Korenmann (1996) provide evidence that obesity has more deleterious effects on the self-esteem of white than black or Hispanic females.

more or less monotonic negative relationship between BMI and earnings predicted from the IV estimates.

Much can be done to clarify the interpretation of our results. Although health expenditures do not appear to drive the patterns, it is unclear whether the findings for women reflect labor market discrimination or some other cause. For example, females working in occupations requiring physical interaction might be subject to particular physical scrutiny. Adding controls for broad occupational categories slightly reduces the gradient of the wage function for females, consistent with occupational sorting; however, definitive answers to this question require controlling for occupational categories measuring the level of public interaction. Some results, particularly for males, are sensitive to the choice of specifications and we poorly understand why the results differ so starkly for whites and blacks. It would also be desirable to model medical expenditures simultaneously with earnings, using data from a single source, to get a better sense of the extent to which employers trade-off wages for health expenditures.

These caveats notwithstanding, our analysis provides useful guidance for interpreting prior studies and conducting future research. First, when examining how BMI is related to earnings (and probably other outcomes), it is important to allow for a variety of possible patterns rather than initially assuming that obesity is "where the action is." Indeed, we find little evidence of an "obesity penalty" per se but instead often show that the conditional wage function is maximized at low levels of BMI, where excess weight is almost certainly not a key factor. Although we suspect that our results provide evidence of "beauty" or "appearance" effects, additional examination of these possibilities is needed. Second, the relationships are often highly non-linear and benefit from models that permit considerable flexibility. We obtain this using our semi-parametric specifications but at the cost of considerable computational complexity. Simpler, although somewhat less flexible, modeling techniques might involve the use of higher order polynomials or linear splines. One possibility is to employ univariate non-parametric methods (without controls other than body weight) to establish the basic pattern, which then guide the choice of parametric models containing the full set of covariates.

Tables and Figures

Table 1.1: Simulated Wage Differences Using Semiparametric Models

Wage Difference ($) Relative to Predicted Earnings at Reference BMI										
	Women (reference BMI = 23)					**Men (reference BMI=27)**				
BMI	20	25	30	35	40	20	25	30	35	40
Full Sample										
SPM	-.93*	-.96**	-1.51***	-2.62***	-2.89***	-3.19**	-.33	-.21	-.81	-1.72
	(.351)	(.277)	(.317)	(.311)	(.407)	(.928)	(.172)	(.195)	(.524)	(.993)
SPM-IV	-.24	-1.13*	-2.19***	-2.17**	-2.17**	2.28	1.84	-2.34***	-5.70***	-10.05***
	(.548)	(.319)	(.477)	(.658)	(.790)	(1.755)	(.574)	(.631)	(1.025)	(1.82)
Whites										
SPM	-.90†	-1.02*	-1.93***	-3.51***	-3.50***	-3.87*	.01	-1.31*	-1.89*	-2.94†
	(.481)	(.419)	(.464)	(.717)	(.718)	(1.572)	(.668)	(.652)	(1.525)	(1.52)
SPM-IV	-.84	-1.45†	-1.79*	-1.17	.000	2.07	2.14**	-2.82***	-6.81***	-11.17**
	(.841)	(.686)	(.791)	(.852)	(1.274)	(2.101)	(.550)	(.565)	(1.29)	(2.660)
Blacks										
SPM	-.49	.05	-.315	-1.16*	-1.17*	-2.82**	-.68**	.60*	.507	-.43
	(.355)	(.140)	(.316)	(.437)	(.437)	(1.02)	(.18)	(.210)	(.599)	(1.06)
SPM-IV	.02	-.27	-1.33**	-2.15**	-2.63**	3.61	-.18	-.25	-5.09**	-8.63**
	(.498)	(.223)	(.446)	(.619)	(.777)	(2.324)	(.559)	(.847)	(1.101)	(1.76)

Note: Results from Semi-Parametric (SPM) Models; Standard Errors in Parenthesis. ***p<.001, **p<.01, *p<.05 †p<.10

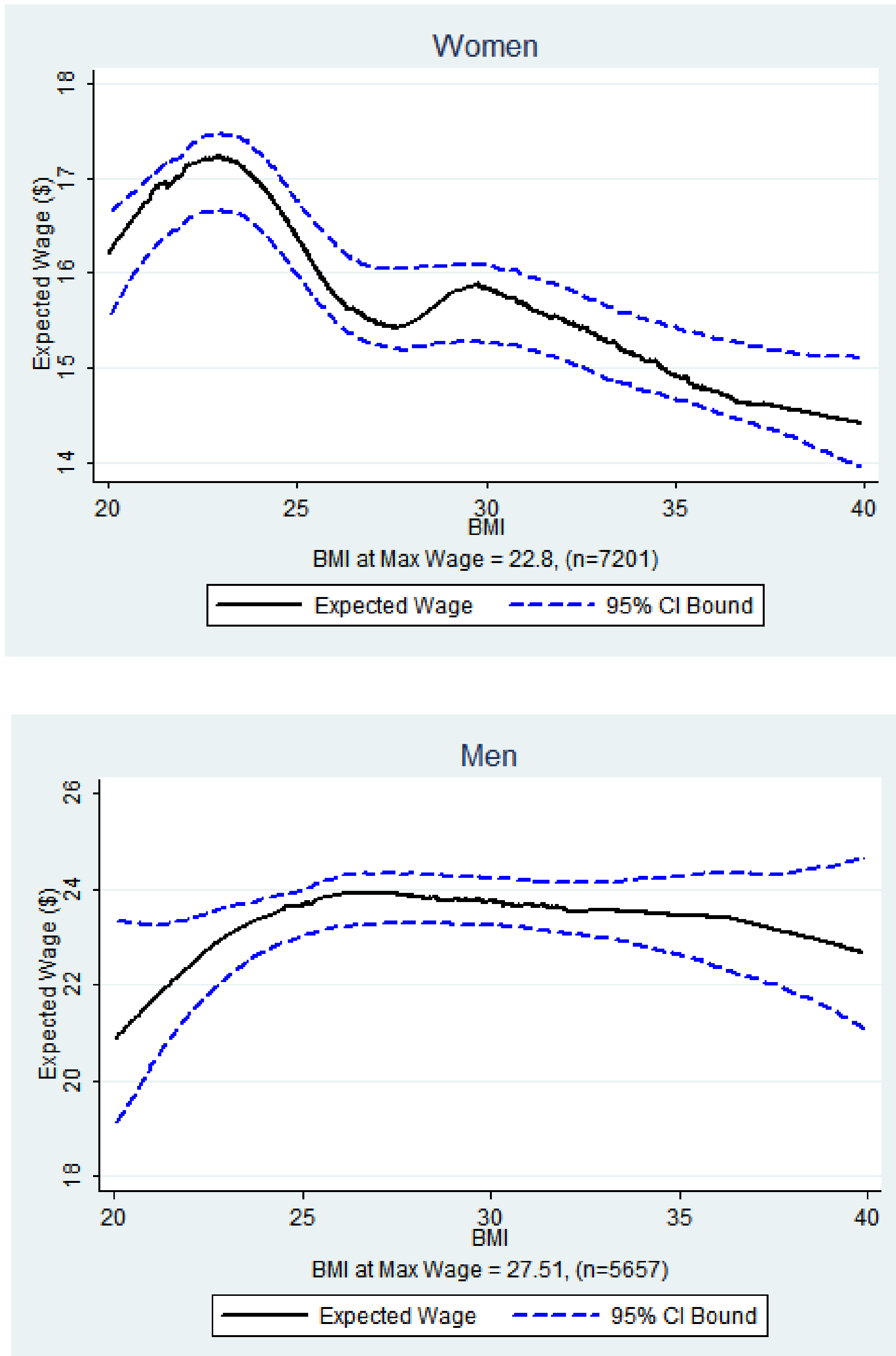


Figure 1.1: BMI and Expected Wages, Full Sample

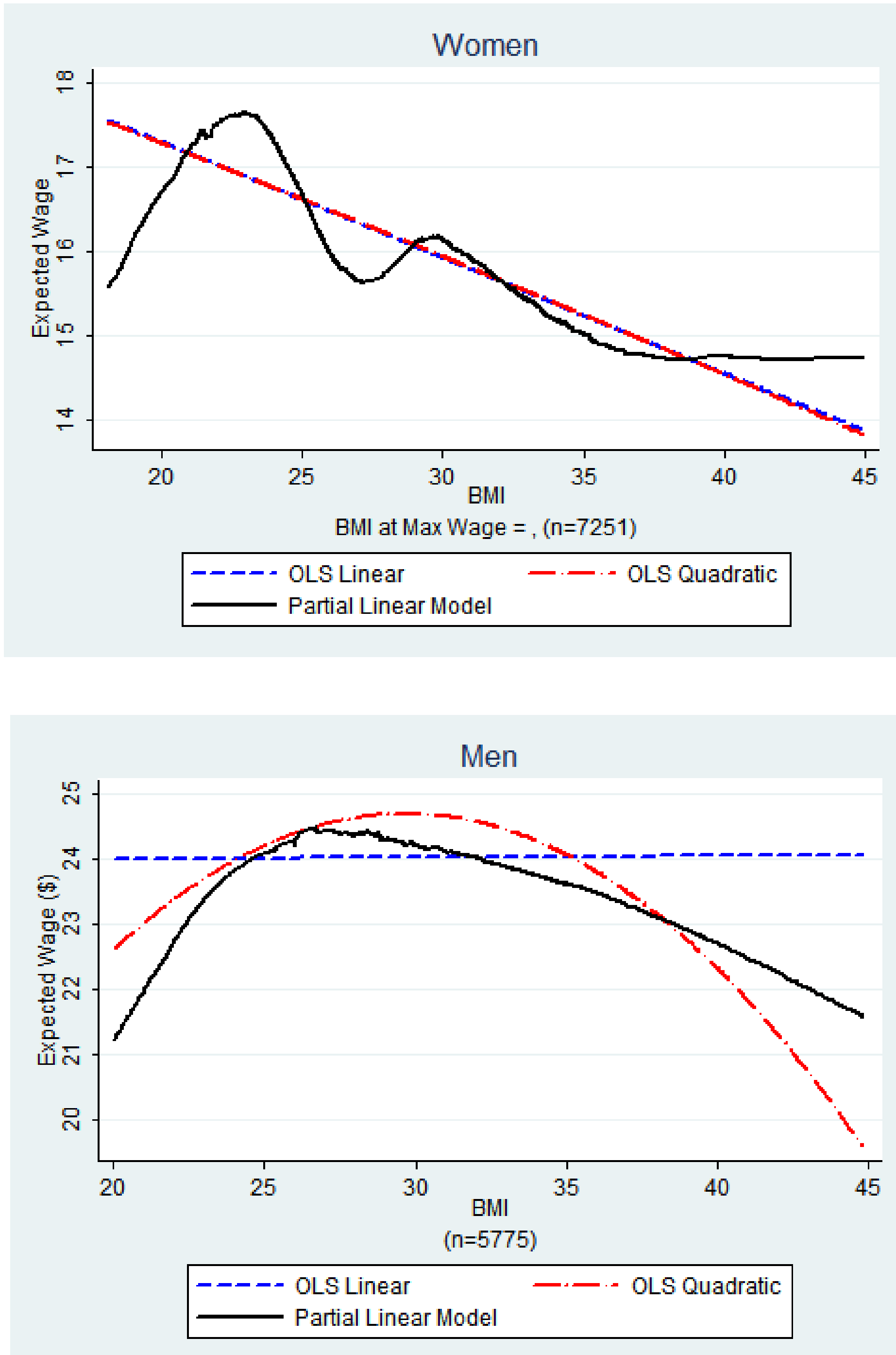


Figure 1.2: Comparison of Three Estimation Models

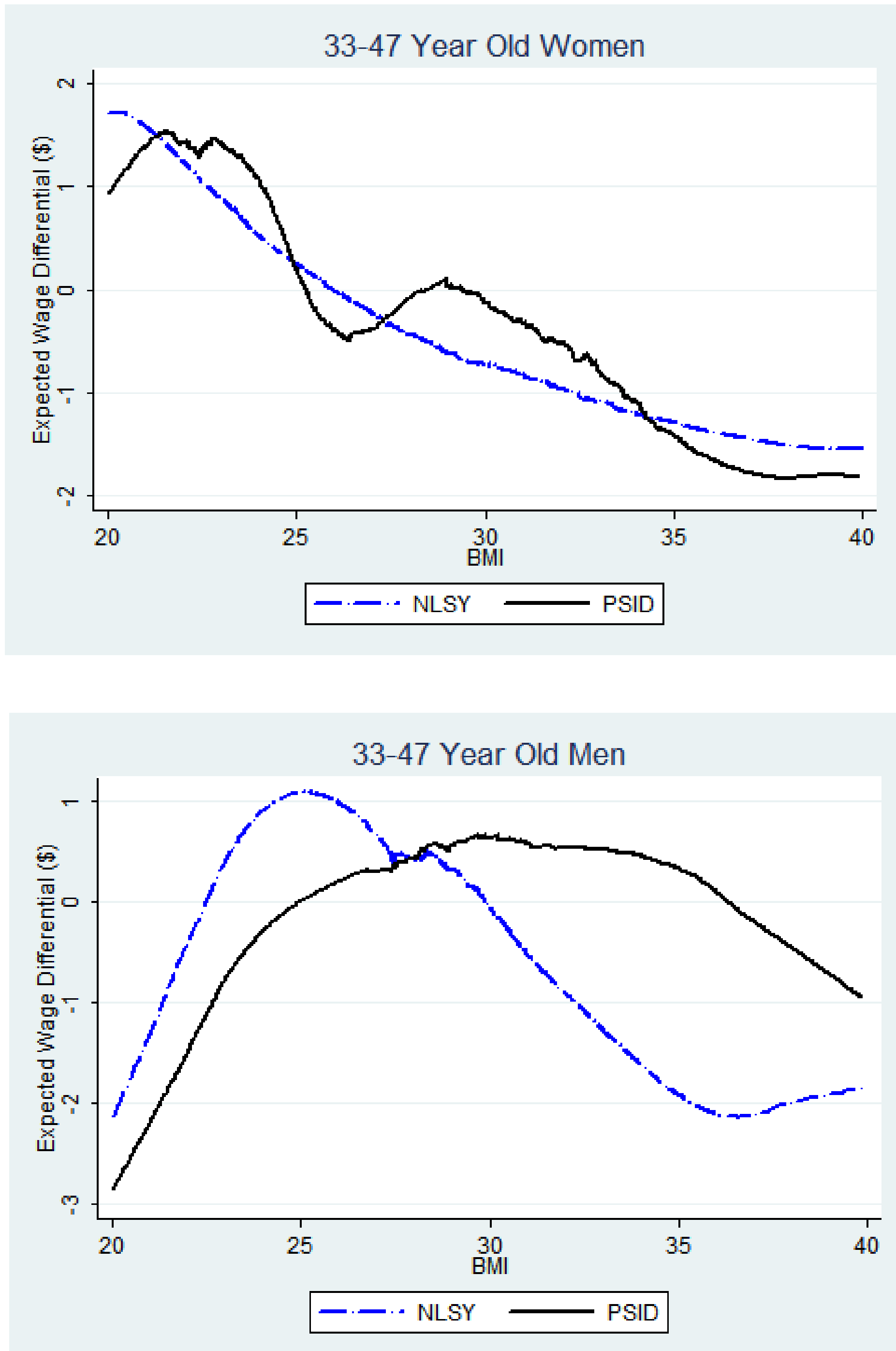


Figure 1.3: BMI and Estimated Wage Differentials, PSID - NLSY Comparisons

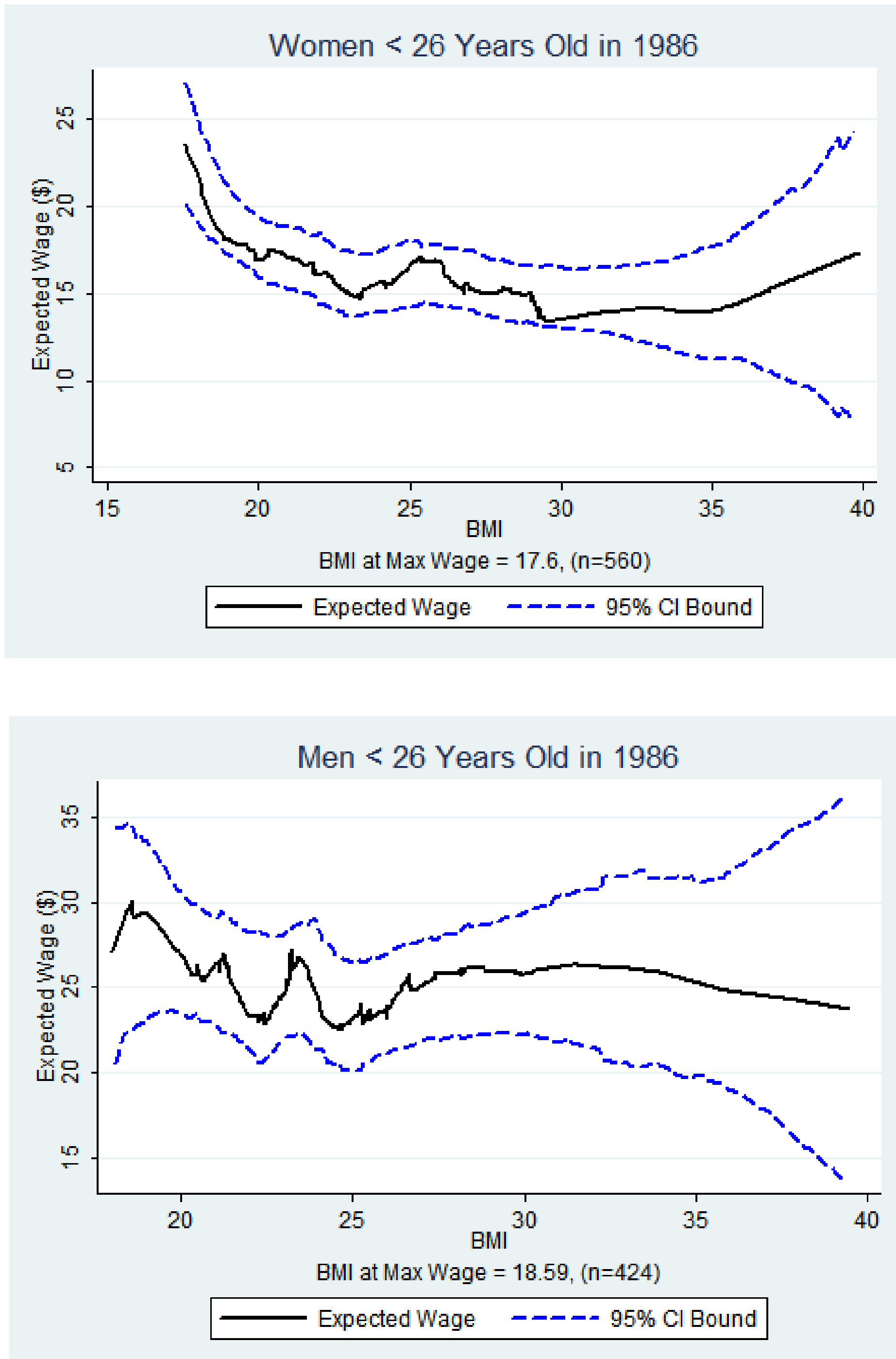


Figure 1.4: Lagged BMI and Expected Wages

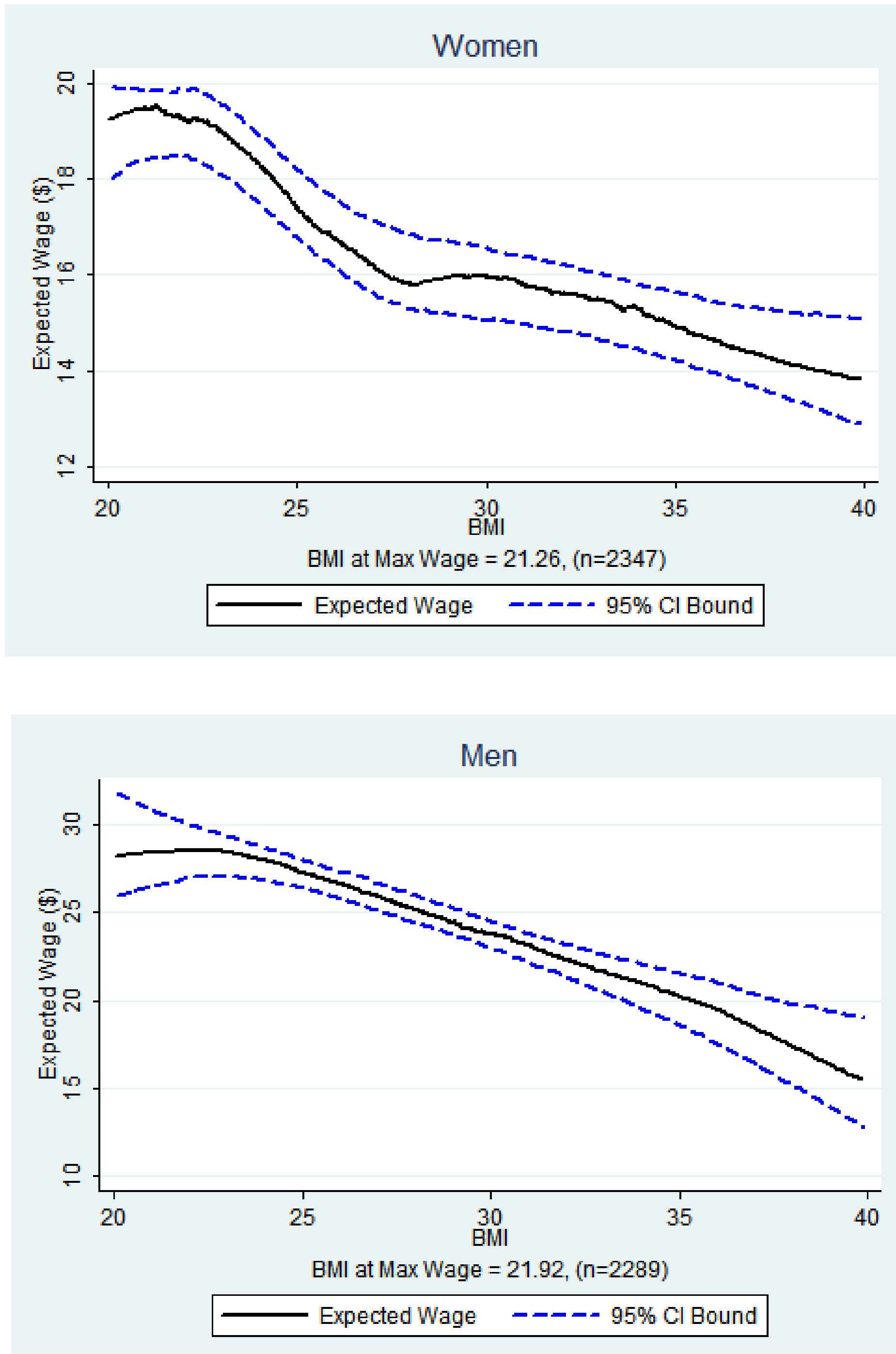
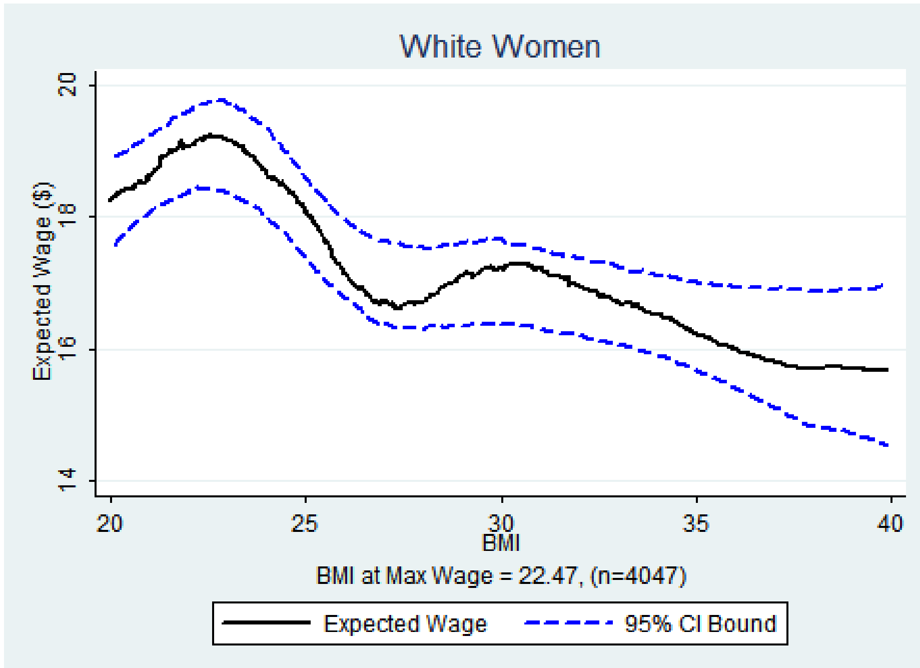


Figure 1.5: Instrumental Variables Estimates



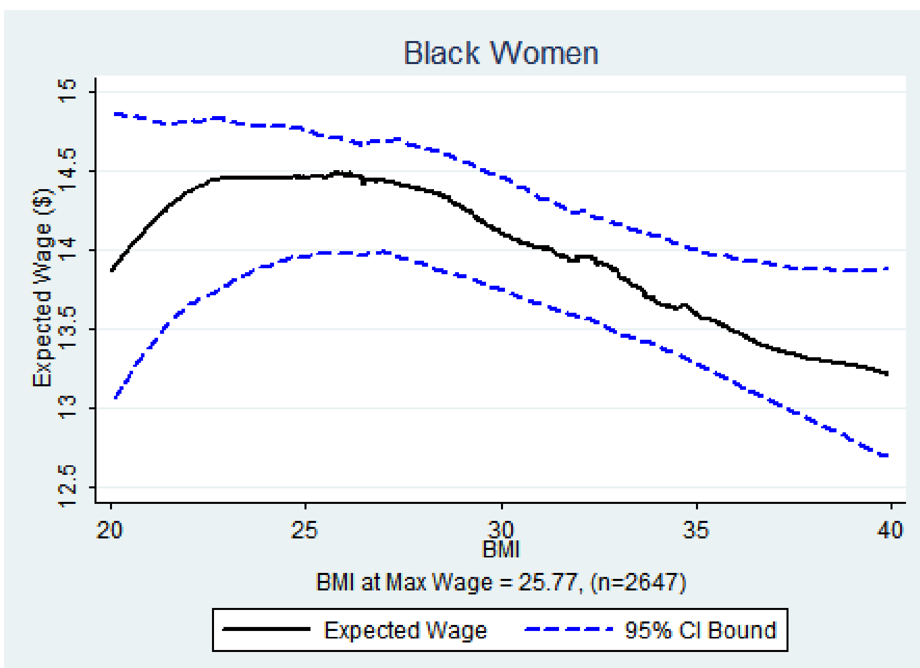


Figure 1.6: BMI and Expected Wages of Women, by Race

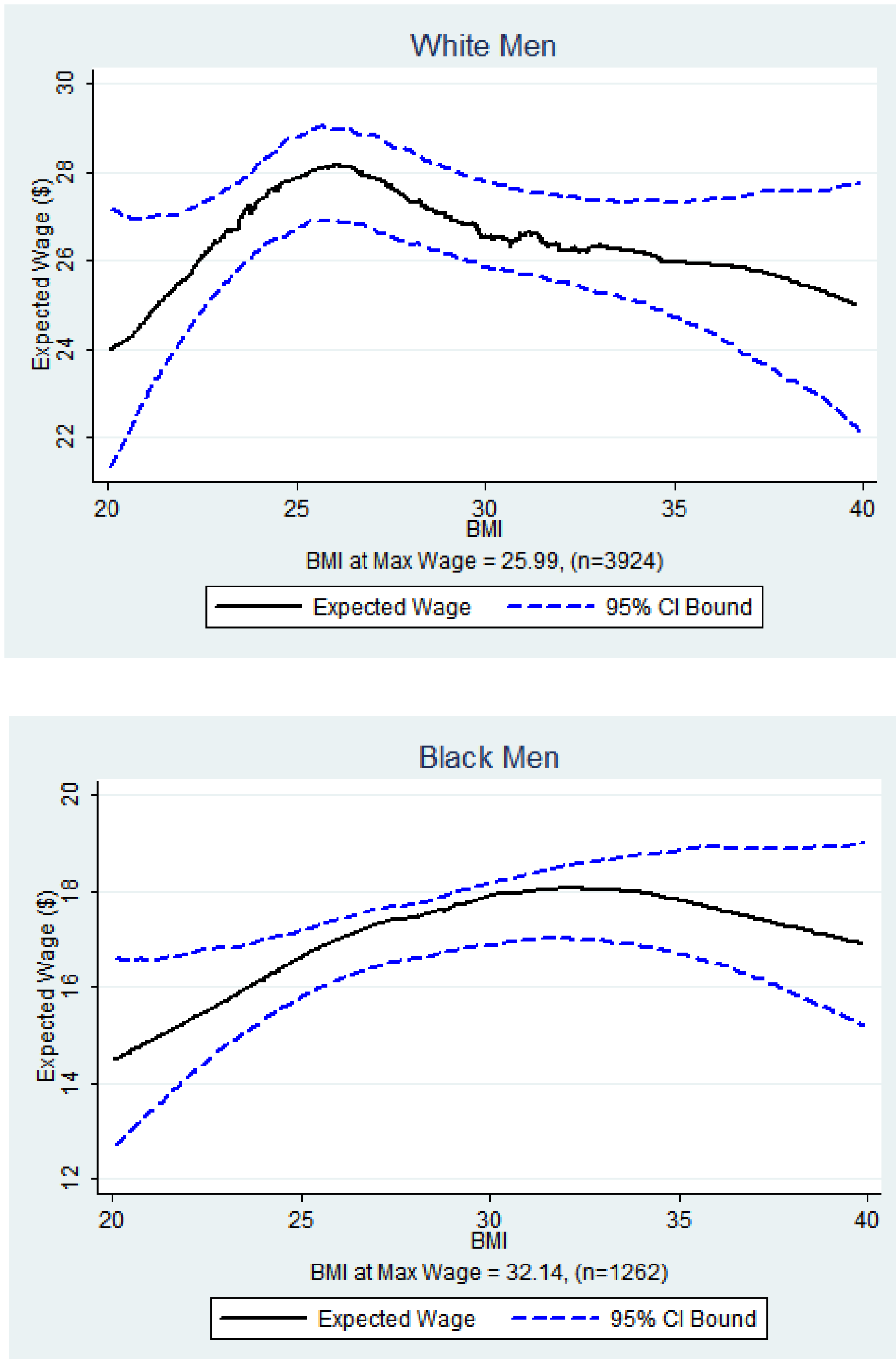


Figure 1.7: BMI and Expected Wages of Men, by Race

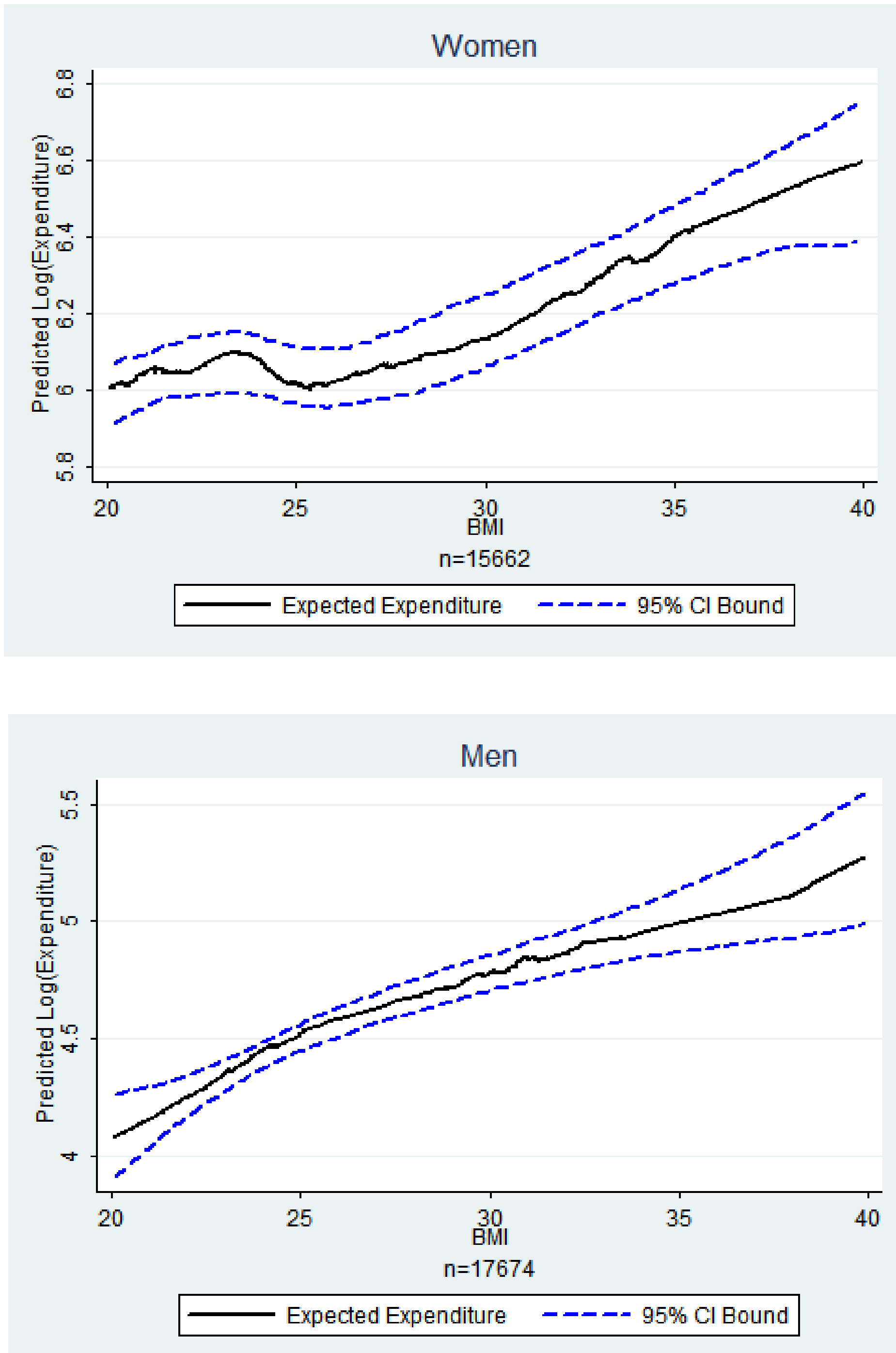


Figure 1.8: BMI and Expected Medical Care Expenditures

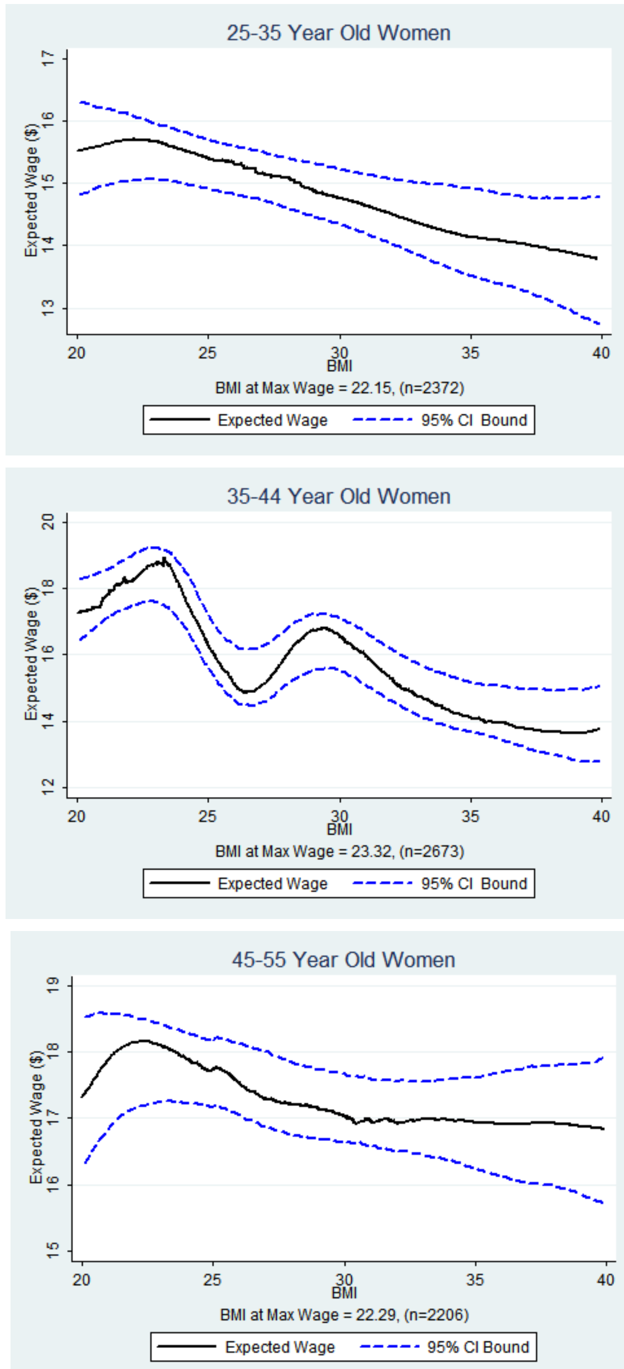


Figure 1.9: BMI and Expected Wages, Females by Age

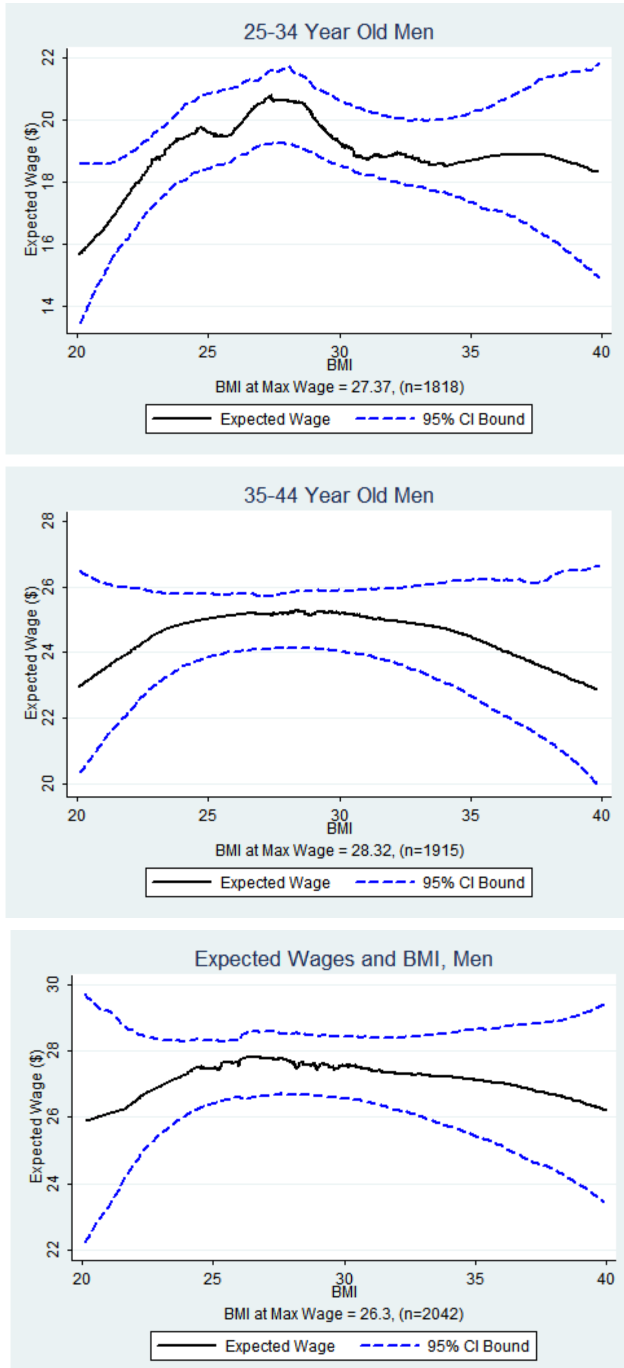


Figure 1.10: BMI and Expected Wages, Males by Age

Appendix: Non-parametric Smoothing Methods and Additional Econometric Estimates

Kernel regression drops the assumption of linearity and models the expectation of the dependent variable as a weighted mean at every point in the distribution of the independent variable. For example, the oft-used Nadarya-Watson kernel estimator can be defined as

$$\hat{r}_n(x) = \sum_{i=1}^{n} \ell_i(x) Y_i \tag{1.2}$$

where $\hat{r}_n(x)$ is the predicted value of y at a given value x, and the weights are defined by the kernel function:

$$K(x) = \frac{70}{81}(1 - |x|^3)^3 I(x) \tag{1.3}$$

where

$$I = \begin{cases} 1 & \text{if } |x| \leq 0 \\ 0 & \text{otherwise.} \end{cases}$$

The choice of the kernel function–Gaussian, uniform, Epanechnikov–generally does not affect the result. The weighting function, $\ell(x)$ is defined as

$$\ell_i(x) = \frac{K(\frac{x - x_i}{h})}{\sum_{j=1}^{n} K(\frac{x - x_i}{h})} \tag{1.4}$$

where h is the bandwidth or smoothing parameter. This kind of estimator has the advantage of allowing for highly non-linear relationships that are frequently missed even with linear estimators that include quadratic, cubic, and higher order terms.

In our analysis, we use local linear regression, which is similar in spirit to kernel regression, but instead of modeling the data with a locally weighted average, it uses a locally weighted linear regression. Local linear regression relaxes the linearity assumption of OLS and minimizes both boundary bias and design bias introduced by the kernel framework.[33] In general, we define the estimator and kernel as in equation 2.6, but define $\ell(x)$, X_x, and W_x as follows.

$$\ell(x) = e_1^T (X_x^T W_x X_x)^{-1} X_x^T W_x$$

$$e_1 = (1, 0, 0, ...)^T$$

$$X_x = \begin{bmatrix} 1 & x_1 - x \\ 1 & x_2 - x \\ 1 & x_3 - x \\ \vdots & \vdots \\ 1 & x_n - x \end{bmatrix}$$

$$W_x = \begin{bmatrix} w_1(x) & 0 & \cdots & 0 \\ 0 & w_2(x) & & \vdots \\ \vdots & \cdots & \ddots & \\ 0 & \cdots & \cdots & w_n(x) \end{bmatrix}$$

$$w_i(x) = K(\frac{x - x_i}{h}) \tag{1.5}$$

This formulation implies that the predicted value for a given value of x is the inner product of the first row of $\ell(x)$ with Y.

The choice of smoothing parameter, h, involves the tradeoff between bias and variance, as h defines the window of observations that will be used in local regression. For non-linear functions, small windows of observations give high variance and low bias, whereas large windows offer the converse. We choose the bandwidth by selecting the span, k, the fraction of the data to include in the linear estimate, to minimize mean squared error ($bias^2 + variance$) for the estimator. This implies that for each realization of x the bandwidth changes according to the distance to the observation $(k * N)/2$ observations away. In

[33] On this point, see Wasserman (2006), 73ff., Fan and Gijbels, pp.17-18, 60ff.

particular, we minimize the leave-one-out cross-validation score over the range of the span. The cross validation score is defined as

$$CV(k) = \frac{1}{n}\sum_{i=1}^{n}(Y_i - \hat{r}_{(-i)}(x_i))^2 \tag{1.6}$$

where $\hat{r}_{(-i)}$ is the estimator derived from leaving out the the i^{th} observation.[34]

[34]When smoothing the dependent variables, we execute least-squares cross validation at the roughly 500 points .2 percentile points apart in the middle 95 percent of the distribution of BMI.

Table 1.2: Semi-Parametric Regression Results for Women

	Full Sample	Whites	Blacks	Age<26 in 1986	IV
Black	-1.149***			-1.597	-1.986**
	(0.269)			(0.978)	(0.766)
Hispanic	-2.865***				3.532*
	(0.537)				(1.588)
Age	0.054***	0.106***	-0.005	-0.404*	0.032
	(0.014)	(0.020)	(0.021)	(0.190)	(0.031)
Year 2001	0.214	0.298	0.151	0.989	-0.059
	(0.340)	(0.474)	(0.498)	(1.168)	(0.696)
Year 2003	1.184***	1.323**	0.988*	3.173*	0.318
	(0.332)	(0.462)	(0.497)	(1.352)	(0.673)
Year 2005	0.585*	0.660	0.104	2.866*	0.518
	(0.271)	(0.409)	(0.352)	(1.414)	(0.527)
Number of Kids	-0.055	0.317*	-0.269*	-0.540	-0.025
	(0.099)	(0.151)	(0.130)	(0.368)	(0.193)
Married	0.638*	0.174	0.708*	-0.156	0.534
	(0.261)	(0.397)	(0.349)	(0.905)	(0.489)
Child Under 2	2.253***	3.472***	0.661	2.429	2.263**
	(0.381)	(0.570)	(0.502)	(1.389)	(0.734)
Northeast	3.284***	2.313***	4.768***	5.040***	3.506***
	(0.343)	(0.462)	(0.570)	(1.329)	(0.651)
Midwest	0.382	-0.108	0.795	-0.154	1.320*
	(0.283)	(0.391)	(0.411)	(1.089)	(0.552)
West	2.372***	1.817***	3.893***	1.246	3.361***
	(0.338)	(0.455)	(0.679)	(1.166)	(0.667)
HS Dropout	-3.008***	-3.482***	-1.786***	-2.624*	-3.597***
	(0.364)	(0.621)	(0.447)	(1.158)	(0.784)
Some College	1.386***	1.367***	1.520***	3.430***	0.350
	(0.269)	(0.395)	(0.352)	(0.942)	(0.531)
College Graduate	7.677***	7.235***	8.140***	11.803***	7.266***
	(0.297)	(0.396)	(0.477)	(1.517)	(0.594)
Job Tenure (Mos)	0.024***	0.025***	0.025***	0.034***	0.027***
	(0.001)	(0.002)	(0.002)	(0.006)	(0.003)
IV Residual					-0.015
					(0.162)
Constant	0.010	0.069	0.043	-0.111	0.014
	(0.107)	(0.155)	(0.147)	(0.354)	(0.206)
N	7251	4047	2638	544	2369

Note: Regression coefficients for supplementary covariates. Standard errors in parenthesis. ***p<.001, **p<.01, *p<.05

Table 1.3: Semi-Parametric Regression Results for Men

	Full Sample	Whites	Blacks	Age<26 in 1986	IV
Black	-5.310***			-3.833*	-5.171***
	(0.580)			(1.824)	(1.143)
Hispanic	-7.346***				-1.536
	(0.968)				(3.082)
Age	0.240***	0.303***	0.085*	-0.102	0.375***
	(0.028)	(0.037)	(0.041)	(0.379)	(0.051)
Year 2001	0.570	0.756	-0.258	0.814	1.602
	(0.635)	(0.825)	(0.936)	(2.124)	(1.140)
Year 2003	0.662	0.846	-0.470	-1.708	1.320
	(0.625)	(0.822)	(0.902)	(2.538)	(1.141)
Year 2005	0.541	0.844	-0.896	0.559	0.618
	(0.550)	(0.745)	(0.718)	(2.919)	(0.919)
Number of Kids	1.142***	1.945***	-0.317	-0.719	1.783***
	(0.206)	(0.289)	(0.282)	(0.691)	(0.368)
Married	2.641***	3.222***	2.613***	4.011*	3.501***
	(0.561)	(0.788)	(0.699)	(1.768)	(0.973)
Child Under 2	0.482	0.330	-0.466	8.792***	-0.075
	(0.762)	(1.051)	(1.063)	(2.646)	(1.296)
Northeast	4.951***	5.772***	2.360*	5.169	5.590***
	(0.660)	(0.838)	(1.122)	(2.648)	(1.135)
Midwest	0.776	0.769	1.534	1.495	0.698
	(0.561)	(0.734)	(0.791)	(1.794)	(1.028)
West	1.335*	1.795*	2.325*	-1.126	1.254
	(0.614)	(0.820)	(1.045)	(2.216)	(1.120)
HS Dropout	-3.942***	-4.056***	-2.050*	-6.972**	-5.171***
	(0.739)	(1.147)	(0.875)	(2.686)	(1.357)
Some College	3.286***	3.380***	3.498***	4.101*	3.371***
	(0.568)	(0.761)	(0.726)	(1.920)	(0.962)
College Graduate	11.720***	12.349***	7.031***	26.354***	11.987***
	(0.549)	(0.703)	(0.872)	(2.325)	(1.132)
Job Tenure (Mos)	0.013***	0.009**	0.021***	0.020*	0.001
	(0.002)	(0.003)	(0.003)	(0.009)	(0.004)
IV Residual					0.735
					(0.434)
Constant	0.322	0.037	0.341	-0.259	-0.004
	(0.212)	(0.282)	(0.291)	(0.676)	(0.360)
N	5775	3924	1262	427	2333

Note: Table shows regression coefficients for supplementary covariates. Standard errors in parenthesis. ***p<.001, **p<.01, *p<.05

CHAPTER II

BMI, WAGES, and AGE

Abstract

Previous research generally finds that obesity negatively affects wages for women and does not affect wages for men. But this literature has for the most part focused on young workers and has not examined whether the effect of obesity might change as people age. In this essay, I examine the effect of obesity–and body mass more generally–on wages across the age distribution, using conventional parametric and more flexible semiparametric approaches. My parametric results suggest that the literature may overstate the effect of BMI and obesity on wages for women and almost certainly understates any negative association for men. For women, my results show that the negative effects of BMI and obesity are concentrated among women between 25 and 35 years old. While women in this age group experience an average 0.5 to 0.7 percent decrease in wages for each point increase in body mass (roughly 7.5 pounds), women over 40 will suffer a 0.25 percent decrease in wages for each extra point of body mass, and may not experience any wage penalty at all. Similarly, women who are 31-35 years old experience a 7.7 percent decrease in wages for being obese, while women over 40 experience only a 3.9 percent decrease. More flexible models largely confirm these results. For men, my parametric results indicate that, for those who are in their 20's or early 30's, BMI has no effect on wages or is associated with a small increase; this is consonant with the rest of the literature. However, for men over 35, the effect of extra body mass is clearly negative: an extra BMI point brings with it a 0.3 percent decrease in wages for men 36-40 years old, and for men over 40, an extra BMI point is associated with a 0.5 percent decrease in wages. More flexible semiparametric models suggest that the negative

association of BMI and wages may be as much as three times more than these estimates in some ranges of BMI. For instance, these models suggest that men 36-40 years old who have a BMI between 27 and 37 experience a 0.9 - 1.2 percent decrease in wages for each extra BMI point, as opposed to a .3 percent decrease predicted by the linear model.

Introduction

Studies of the correlation between obesity and wages for subjects in the United States have fairly consistently found that obese women earn less than non-obese women; estimates of this negative association range from 1 to 12 percent. Work in this area has also found that obese men earn only slightly less–and may earn slightly more–than their non-obese counterparts (Register and Williams, 1990; Pagan and Davila, 1997; Harper, 2000; Averett and Korenmann, 1996; Sargent and Blanchflower, 1994; Gortmaker et al., 1993; Loh, 1993; Baum and Ford, 2004; Cawley, 2004; Norton and Han, 2008). However, most of this literature examines people who were relatively young when wage data were obtained. Sargent and Blanchflower (1994), for example, use wage data from subjects who were 23 years old; Harper (2000) looks at persons who were 33 at the time of survey; Loh (1993) and Register and Williams (1990) use data on those whose mean age was just over 22 years old; Pagan and Davila (1997) use data for respondents whose average age was 28; similarly, subjects examined by Averett and Korenmann (1996) and Gortmaker et al. (1993) were 27 years old on average. Norton and Han (2008) looked at data for persons who were younger than 27 years at the time of data collection. A few recent studies have included persons over 35 years old: Baum and Ford (2004) use the National Longitudinal Survey of Youth 1979 (NLSY79) through 1998, at which point the age range was 33-41 years, and Cawley (2004) uses data from the NLSY79 through 2000, which includes persons two years older. In addition, Gregory and Ruhm (2009) examine the effect of body mass index (BMI) on wages for persons between 25 and 55 years old, and Han et al. (2009b) examine workers up to 41 years old.[1] Although these studies do include older workers, except for the Gregory and Ruhm

[1] Body Mass Index (BMI) is defined as weight in kilograms divided by height in meters squared.

(2009) and Han et al. (2009b), none has examined how the association between wages and obesity (or BMI) changes as people age. And both Gregory and Ruhm (2009) and Han et al. (2009b) examine this question in supplementary specifications or robustnest checks, not systematically.

There are at least four mechanisms–all unobserved–that might explain a causal relationship between obesity and wages: labor market discrimination, preferences of consumers, marginal productivity differences, and health costs. Although the strength of each of these effects can be expected to change with age, it is unclear whether the total effect of aging would be to reduce or increase the measured effect of obesity on wages. Consider the case when the cause is labor market discrimination. Empirical work by Biddle and Hamermesh (1994) has shown that it is plausible that what we observe as an obesity effect is really an effect of personal appearance–at least insofar as being obese is considered a detriment to physical attractiveness. Although we do not know if the effect of looks on wage changes with age, it seems reasonable to think that the marginal effect of body weight on attractiveness would diminish with age, as the fraction of overweight or obese persons increases.[2] If that were true, then we would expect the obesity wage penalty to decrease with age. Similarly, if the cause were preferences of consumers or the public for persons with low BMI, one might expect that social norms about BMI would be relaxed as persons aged, and thus that the penalty to having a high body mass would decrease. On the other hand, if wages are affected by diminished marginal productivity due to the co-morbidities of obesity, then we should expect the effect of obesity to increase with age, as we know that age-specific prevalence of conditions such as diabetes and dyslipidemia is significantly higher for obese persons as they get older (Finkelstein et al., 2007). Similarly, we would expect the effect of obesity to increase if, as Bhattacharya and Bundorf (2005) argue, employers are able to trade off wages for expected health costs due to obesity, which we know increase with age (Finkelstein et al., 2007).

In this essay, I examine the effect of obesity–and body mass more generally–on wages

[2]Results of a robustness check reported by Gregory and Ruhm (2009) are consistent with this thesis.

across the age distribution, using conventional parametric and more flexible semiparametric approaches. My parametric results suggest that the literature may overstate the effect of BMI and obesity on wages for women and almost certainly understates any negative association for men. For women, my results show that the negative effects of BMI and obesity are concentrated among women between 25 and 35 years old. While women in this age group experience an average 0.5 to 0.7 percent decrease in wages for each point increase in body mass (roughly 7.5 pounds), women over 40 will suffer a 0.25 percent decrease in wages for each extra point of body mass, and may not experience any wage penalty at all. Similarly, women who are 31-35 years old experience a 7.7 percent decrease in wages for being obese, while women over 40 experience only a 3.9 percent decrease. More flexible models largely confirm these results. For men, my parametric results indicate that, for those who are in their 20's or early 30's, BMI has no effect on wages or is associated with a small increase; this is consonant with the rest of the literature. However, for men over 35, the effect of extra body mass is clearly negative: an extra BMI point brings with it a 0.3 percent decrease in wages for men 36-40 years old, and for men over 40, an extra BMI point is associated with a 0.5 percent decrease in wages. More flexible semiparametric models suggest that the negative association of BMI and wages may be as much as three times more than these estimates in some ranges of BMI. For instance, these models suggest that men 36-40 years old who have a BMI between 27 and 37 experience a 0.9 - 1.2 percent decrease in wages for each extra BMI point.

Data

The data used for this analysis come from the NLSY79. The NLSY79 is a survey that was administered yearly from 1979 to 1994 and has been given biennially after 1994; it collects data on a rich set of labor market outcomes, demographics, and, in select years, height and weight for persons born between 1957 and 1965. In 1979, the NLSY was comprised of a sample of 12,686 persons.

Because it has collected information on a wide variety of family background, demographic, and labor market variables over almost three decades, the NLSY79 has been the primary resource for research into the effect of obesity on wages for subjects in the United States.[3] One limitation of the dataset is that it is restricted to a single cohort only now reaching middle age. This restricts the age range over which I can examine the effects of body mass on earnings. Moreover, it is not possible with this data to disentangle cohort from age and year effects. However, the NLSY79 has a far longer panel of height and weight (and, therefore, BMI) information on sample respondents than, for example, the Panel Study of Income Dynamics (PSID), at least for the age cohort that it does represent. This lends statistical power to estimates and allows for a finer examination of the age distribution than the PSID. In addition, results using the PSID, a panel not limited to a single cohort, are qualitatively similar for to those from the NLSY79, which suggests that cohort effects are not driving the results (Gregory and Ruhm, 2009).

I examine hourly wages in this study because, in the context of competitive labor markets, they are a measure of marginal productivity; unlike earnings, they do not involve decisions about hours of work, which may be endogenous to body mass; unlike other kinds of compensation such as health benefits and flex-time, they are easily observed. The hourly wage information that I use is constructed in the NLSY79 through a series of questions about the time unit and rate of pay for the respondent's current or most recent job. For any respondent who reports a time unit rate of pay other than hourly, an hourly rate is constructed by the Center for Human Resources Research (CHRR). The sample includes observations for respondents who were unemployed or out of the labor force at the time of interview, but have had a job since the last interview; in these cases, the wages reported are for the most recent job since the last interview. I include only those over 20 years old and exclude from analysis pregnant women, persons in the military, and self-employed persons and persons with BMI values in the upper or lower one-half of one percent of the BMI distribution. (See discussion below.) Hourly wages are deflated to 2006 basis using

[3] Gregory and Ruhm (2009) is the first study of which I am aware to use another source, the Panel Study of Income Dynamics (PSID) to study the effect of BMI on wages for U.S. subjects.

the CPI. For the estimates shown here, I have changed any wage value that is less than half of the minimum wage to the maximum of the mean value of non-missing wage values for that person and half the minimum wage; additionally, I have changed any wage value in the top one-tenth of one percent of the wage distribution to the minimum of the mean value of observed wages for that person and the 99.9^{th} percentile value.[4]

I construct BMI from self-reported height and weight measures. There are 16 survey years in which weight data are recorded: 1981, 1982, 1985, 1986, 1988, 1989, 1990, 1992, 1993, 1994, 1996, 1998, 2000, 2002, 2004, and 2006. Height data are collected in 1981, 1982, and 1985. Since all respondents have reached the age of 20 by 1985, this is assumed to be their adult stature.

Self-reported height and weight data are known to be subject to reporting error. I have used data from the National Health and Nutrition Survey (NHANES) 1988-94, 1999-2000, 2001-2002, and 2003-2004 to correct for self-reporting bias using a procedure similar to that suggested by Lee and Sepanski (1995), Cawley (2004) and Lakdawalla and Philipson (2007), among others. This procedure takes advantage of the fact that NHANES has both measured and self-reported BMI data; using this procedure, I regress measured height (weight) on self-reported height (weight), their squares, age and its square; I use the coefficients–stratified by gender and race–to predict actual from reported height and weight in the NLSY sample. I use these corrected measures throughout this study.

An emerging strand of the literature on weight and wages has been critical of using BMI as a measure of body fat, preferring other measures such as biolectrial impedence, skinfold thickness, or waist circumference. (Johannson et al., 2009; Wada and Tekin, 2007; Burkhauser and Cawley, 2008b). While it is true that BMI is an imperfect measure of body fat, it has yet to be demonstrated that it is inferior to these other measures by some standard

[4]In all cases, extreme wage values appear to be due to reporting errors. I am sensitive to the argument made by Bollinger and Chandra (2005), who recommend against winsorizing procedures like these. However, whereas the values chosen to replace outliers in Bollinger and Chandra (2005) are dependent on full sample distributions, I choose replacement values here based on the person-level distributions of wages. I have also estimated all of the models shown below using data in which I trimmed only a handful of wage observations–i.e. those above $30,000 per hour. The parametric results were the same; the semiparametric estimates were only marginally affected, mostly in the sense that this trimming exacerbated the well-known characteristic of linear smoothers to overfit (undersmooth) the data.

measure, such as densitometry. As Deurenberg et al. (1991) have noted, the correlation between body fat measured by densitometry and BMI is fairly high–between .78 and .92 for men and women respectively in the age group relevant to this study. Most importantly, almost all of the literature that looks at the relationship between body composition and wages uses BMI, and it is the only measure available in the NLSY79.

The sample design of the NLSY79 affords considerable latitude to researchers interested in labor market outcomes, as it is comprised of three subsamples: a representative cross sectional subsample (6,185 people in 1979), supplementary subsamples (5,295 people in 1979), and a military subsample (1,206 people in 1979). The supplementary subsamples include over-samples of black, Hispanic, and economically disadvantaged white respondents.[5] There are many applications in which one would prefer to use to use the "full" sample, which includes the representative and supplementary subsamples: for example, one might use this sample to get race-specific estimates for the effect of BMI or obesity, as has been done by Cawley (2004). For this application, I show results derived from the representative sample, although the results for the parametric models are the same for the full sample.[6] I show these results here to retain consistency with the semiparametric estimates, for which the smaller sample sizes are more conducive to more computationally intensive procedures.

By restricting the main analysis data to the representative sample, I drop 105,200 of 202,976 possible person-year observations; of these 105,200 observations, 50,439 have valid wage observations. From the remaining sample, 97,776 person-years, I drop 26,919 person-year observations with missing wage values and a further 2,252 person-year observations that are missing BMI. (12,722 observations are missing both wages and BMI.) I further drop 686 person-year observations of those in the military, 4,568 person-years for those who are self-employed, 1,150 person-years for pregnant women, 4,513 person-years of those under 21, and 377 person-year for those who have BMI's in the top or bottom one-half percent of the BMI distribution.[7] That leaves a sample of 28,182 person-years for women, comprised

[5]The NLSY dropped the military and economically disadvantaged white subsamples in 1985 and 1991, but has revised the sample weights to reflect the new probabilities of sampling.

[6]See Appendix A for further discussion.

[7]The total number of persons in the possible 97,776 person-year representative sample who are in the

of 2,911 persons, and 29,192 person-years for men, comprised of 2,804 persons. The women and men in the sample are both observed for different lengths of time in the survey. Table 2.1 shows the frequencies with which we observe women and men in the NLSY sample. For women, the person-level sample size is somewhat smaller than the one in Gregory and Ruhm (2009), who use a PSID sample comprised of 3,564 women; for men, the person-level sample size is comparable to that used in Gregory and Ruhm (2009), which is comprised of 2,638 men.

As mentioned, there are 2,252 who have valid wage values but are missing height and weight and thus BMI. One way to deal with missing values would be to set the missing value to zero and create a variable indicating that the variable is missing and set it to one, as has been done by some others.(Baum and Ruhm, 2007; Cawley, 2004). In this study, however, it would be inappropriate to set BMI to zero in flexible specifications. For this reason, I have discarded these observations. I show in Appendix B that whether one imputes BMI or discards these observations does not affect the results.

Aside from BMI and age, other controls in the parametric and semiparametric models include race (in OLS models), education, total labor market experience, job tenure, weekly hours, Census region, indicator for SMSA, rural residence, number of children in the household, part-time status (under 20 hours a week), indicator for school enrollment, marital status, and indicators for years of survey. Among these variables, only race and age are not chosen by the subjects of the survey. The rest of the controls represent choices that are arguably related to both body mass and wages in complicated ways. I include them here to make the results of this study comparable to most of the previous literature, which, despite their endogeneity to weight and wages, has for the most part used these controls as a way to isolate the effect of body mass on wages.[8] Means of dependent and explanatory variables with the exception of the year dummies are shown in Table 3.3.

military, self-employed, pregant, under age 21, or who have extreme BMI observations is 2,022, 5,348, 1,660, 7,590, and 611, respectively. The difference between these totals and the marginal number dropped listed in the text is the number who are also missing wages or also missing wages and BMI.

[8]A notable exception is Han et al. (2009a), which estimates the indirect effect of teen BMI on education and education on occupation choice and wages.

Empirical Specifications

As mentioned above, this study addresses the change in the effect of BMI on wages over the age distribution. The general model that I employ to model wages as a function of BMI is

$$w_{it} = X_{it}\beta + [(G_{it})|Age] * \delta + \varepsilon_{it}, \tag{2.1}$$

where w_{it} is the log of hourly wages, X is vector comprised of the variables described above, G is the parameterization of BMI. In parametric models, $G = BMI_{it}$ or $G = W_{it}$, where W is a vector of indicator variables for BMI category: $W_{it} = [BMI_{it} < 18.5, 18.5 \leq BMI_{it} < 25, 25 \leq BMI_{it} < 30, 30 \leq BMI_{it}]$. In all specifications, the estimates of this marginal effect are interacted with age: I use age categories 21-25, 26-30, 31-35, 36-40, and 41 and older.[9]

As is well-known, the mechanisms through which body mass might have its effect on wages are not always observed on an individual level. For example, although body mass can be thought of as an index of underlying health, there is wide variation in the actual levels of health experienced by people with high and low body mass. In addition, overweight or obese persons might have higher preferences for immediate gratification (consumption) over investment (i.e. exercise, dieting), and so might not invest as much as they might otherwise in education or training, which clearly would affect wages. Finally, it is also known that persons with higher body mass are more likely to suffer from depression independent of observed health conditions (Dong et al., 2004). These factors, among others, which are correlated with BMI but not observed will tend to bias estimates of the marginal effect of BMI on wages in an ordinary least squares (OLS) econometric framework. With this in mind, we can restate the error term in (2.1) as

$$\varepsilon_{it} = \mu_i(X_i, BMI_i) + \kappa_{it}(X_{it}, BMI_{it}) + \upsilon_{it}, \tag{2.2}$$

[9] These age groups are large enough to measure effects meaningfully, but not so large as to elide differences in effects at different moments in workers' careers.

where μ represents time-invariant unobserved personal characteristics, X_i and BMI_i represent the history of observables and BMI for person i, κ represents time-varying unobserved factors, and X_{it} and BMI_{it} are the current period observables for person i; υ is a random error.

Although it would be ideal to model μ as well as κ, strategies to identify both of them have been in short supply. One of the ways proposed to model endogenous factors quite generally has been to instrument for own BMI with the BMI of a sibling (Cawley, 2004), parent (Kline and Tobias, 2008), or child (Shimokawa, 2008). Whatever the strengths or weaknesses of this approach, it poses a limiting practical problem for this application. Namely, using this strategy even as a robustness check would make the size of age-specific samples too small to draw meaningful inferences.

A common strategy for isolating μ, which I employ here and has been employed by others (Baum and Ford, 2004; Shimokawa, 2008), is a fixed effects (FE) strategy. In this model, we have

$$w_{it} = X_{it}\beta + [(G_{it})|Age] * \delta + \mu_i + \upsilon_{it}, \tag{2.3}$$

where μ represents time invariant components of personal attributes like health, preferences for consumption, and disposition for depression, among others. There are at least two concerns about this strategy: first, and obviously, it will not capture κ; second, it seems to assert the there is a contemporaneous effect of BMI on wages, which some find implausible.

With respect to the first objection, the literature suggests that FE and sibling IV models will yield qualitatively similar results. For example, although Cawley (2004) and Baum and Ford (2004) prefer instrumental variable (IV) and fixed effect (FE) specifications respectively, both find that women's wages are negatively affected by being obese and that men's wages are less so. It might be objected that κ may be due to reverse causality, which could operate independently of any fixed effects. The literature has addressed this by using lagged BMI as a regressor of interest in OLS (Averett and Korenmann, 1996; Cawley, 2004) or FE models (Shimokawa, 2008). These studies rarely find robust differences between estimates using contemporary and lagged values of BMI. Indeed, I have estimated all of the parametric models using a lagged BMI as the regressor of interest with no change in the results. In addition, Gregory and Ruhm (2009) have used 13-19 year lags of BMI in semiparametric regressions and gotten results that are qualitatively similar to models using contemporaneous BMI. [10] With respect to the second objection, if one views current BMI as embodying the effect of health and capital investments since the end of the previous period, then current period BMI is actually a good regressor of interest; it will be correlated with both recent health investments and current health, independent of permanent health stock captured by μ.

While linear models of the marginal effect of BMI on wages will be informative about variation around the mean value of BMI, recent work (Shimokawa, 2008; Kline and Tobias, 2008; Gregory and Ruhm, 2009) has suggested that more flexible models can highlight features of expected wages over the entire BMI distribution that are missed by linear models. In order to examine whether this holds in this application, in addition to the parametric specifications, I employ semiparametric models that allow all regressors other than BMI to vary linearly with the log of wages, but allow the form of the relationship between BMI and wages to be dictated by the entire BMI distribution, rather than the mean only. Model 2.3 in this context becomes

$$w_{it} = X_{it}\beta + [f(BMI_{it})|Age] + \mu_i + \upsilon_{it}. \tag{2.4}$$

Although this method is more resource intensive, it also can offer a more accurate sense of the complexity of the conditional wage function. For example, Gregory and Ruhm (2009) have shown that the conditional wage distribution for women in the PSID is not particularly well approximated by linear models and that it peaks at a BMI level of about 22.5. Their results are important because they suggest that what has been called an obesity effect is almost certainly not one for women, which would imply that the health effects of body

[10] I used the previous survey wave observation of BMI in these models to preserve sample size in the youngest age groups.

weight are not the cause of decreased wages, at least not alone.

For the flexible estimates, the method I employ is similar to that described by Gregory and Ruhm (2009), Shimokawa (2008), and Yatchew (2003), among others, and is specifically based on the estimator proposed by Li and Racine (2007).[11] In brief, it involves an extension of univariate kernel regression, in which one estimates a weighted regression for each data point in the sample, including in each regression only the observations that fall within a chosen window width (called the bandwidth) of the data point in question. At the extremes, a conventional regression is one in which all data points are included in the window for the regression, and a model that simply interpolates the data is one that includes only one data point in the window width. One chooses the appropriate bandwidth by estimating the local linear regression over many bandwidths and choosing the one that minimizes the mean squared error of the estimate. For applications like the current one, in addition to using this smoothing regression on the variable of interest, one also estimates the effect of other variables in a linear framework. Hence, the semiparametric estimates are also called "partially linear." This method includes purging the dependent variables of the correlation with BMI by use of such a non-parametric smoothing regression and using those purged variables in a linear regression that includes everything except BMI. One uses the residuals of that regression to form the estimate of $\hat{f}(.)$. Fixed effects are established in this context through taking differences in the smoothed variables and regression residuals for each person. I describe the estimation process for the fixed-effect partially linear model (2.4) in more detail in Appendix C.

Descriptive Results

Obesity Wage Differentials

Although I intend to treat BMI more flexibly than has generally been done in the literature, it is useful to begin with a conventional treatment of body mass–particularly, obesity–as

[11] This is itself a variant of the estimator proposed by Robinson (1988).

related to wages. As a simple graphical illustration of the differences in the relationship between obesity and wages across the age distribution, Figures 2.1 and 2.2 show the differences in expected log wages between obese and non-obese persons stratified by age category (21-25, 26-30, 31-35, 36-40, and over 40 years old) for women and men, respectively. The figures also show 95% confidence intervals for these differences. As mentioned above, most studies of the effect of obesity on wages have looked at only the youngest three of these age groups, but the figures both suggest changes in the effect of BMI as people get older.

The results for women, in general, show a "U-shaped" pattern of differences; while obese women earn roughly 14 percent less per hour than the non-obese in the sample when they are 21-25 years old, this unconditional difference increases to 20 percent for women 26-30 or 31-35 years old, and then decreases in magnitude to roughly 14 percent for women over 40. The 95% confidence intervals for these estimates suggest that the differences in all age groups alone are statistically different from zero. (They are different at $p < .01$ as well.) However, differences between age groups (i.e. if we compare differences in one age group to another) are not significant at conventional levels.

Certainly, we should be cautious about generalizing from these results about the effect of body mass on wages. However, it is worth noting that the estimates are consistent with those of Baum and Ford (2004), who also found that the unconditional wage penalty for obese women increased for the NLSY79 sample until 1994, when they were between 29 and 36 years old, and then decreased in 1996 and 1998. Moreover, the average penalty for workers up to 35 years old in my estimates here (18 percent) is roughly the same as that estimated by Baum and Ford (2004) (about 21 percent). However, these figures also highlight that the decrease in the wage penalty has been sustained as women aged. This result is important because it suggests that, to the degree that our current understanding of the wage penalty is based on samples that are heavily weighted toward younger workers, it will offer an incomplete picture of the wage-BMI relationship.[12]

The results for men suggest that the wage penalty increases with age. The difference

[12] I use Baum and Ford (2004) as a comparison because it is the only study which looks at simple mean differences between obese and non-obese persons as they change over time.

in wages for obese and non-obese men is a statistically insignificant 2 percent when they are 21-25 years old. However, for men 26-30 and 31-35 years old, the difference in wages increases to 9 percent and is statistically different from zero at $p < .05$. By the time men are over 35 years old, the unconditional obesity penalty is roughly 13 percent and is statistically different from zero at this level. (All of the wage differences within the age groups are also significant at $p < .01$ with the exception of the difference for those 21-25 years old.) The 95% confidence intervals plotted in the figure show that, if we compare the wage difference for men 21-25 years old with those for men 36-40 and over 40 years old, these difference-in-differences are significant at $p < .05$. If we were to plot 90% confidence intervals, we would see that, in addition, the wage difference for men 31-35 years old is significantly different from that for men 21-25 years old.

This result suggests that the penalty to being obese increases for men as they get older. These results are stronger in terms of magnitude and significance than those for women. They show that the obesity penalty is more than 50 percent greater for men over 40 than for men in their late 20's, and differences between those in their early 20's and men over 35 are statistically significant. While these results for those up to their early 30's are larger than more fully specified regression estimates in the literature–which find little or no penalty for extra body mass–they suggest that for older men being obese may indeed be a detriment to wages.

BMI-Wage Relationships by Age

As mentioned above, recent work has shown that treating body mass in a more flexible manner can improve our understanding of the relationship between body mass and wages. To augment the intuition offered by Figures 2.1 and 2.2, I have also produced univariate non-parametric estimates of log wages as a function of BMI for each age group in Figures 2.3 and 2.4. These figures are scaled to show differences from the mean of log wages for each age group. [13] The figures also use vertical lines to show the first and third quartile of

[13] The figures show the predicted values for the middle 95% of BMI values in the sample.

BMI observations. I include these for two reasons. First, they give us a easy way to gauge how BMI–as opposed to obesity–affects wages on average in these samples. Second, they give us a sense of how changes in the distribution of body mass as people age may alter its effect on wages.[14]

There are a few characteristics of the panels of Figure 2.3 worth highlighting. First, it is not obvious from casual inspection that the marginal effect of BMI decreases all that much as women get older, as is suggested by the results in Figure 2.1. For example, if we compare persons with BMIs of 20 and 30 in the 31-35 and 41 and over age groups, the respective differences in expected log wages are roughly -.16 and -.15. But, second, as is shown by the lines marking the first and third quartiles of the distribution, the middle of the distribution has grown wider and shifted decidedly right for women over 41 years old. The result is that the difference between women in the right tail of the distribution and those more or less in the center has decreased as women on the whole have gotten heavier and as extreme values of BMI have become more common. Although this rightward shift in the distribution of weight for women has been highlighted before (Ruhm, 2007), it has not been remarked in examinations of the effect of BMI or obesity on wages. Finally, the negative effect of BMI on wages for women is not an obesity effect. For all age cohorts, the peak of the wage function is nowhere close to a BMI of 30; rather it lies around a BMI of 22. Although these are unconditional regressions, they nonetheless suggest that we should proceed with caution when using or interpreting models that assume that the wage penalty is only experienced by the obese.

As for the men's results, shown in Figure 2.4, it is not obvious that the effect of BMI on wages is increasing as men get beyond 35 years old. If we just examine the portion of the wage curve that slopes downward (i.e. past the peak) in each of the panels, there is not much change in that part of the wage distribution. In fact, for men 36-40 the downward sloping portion of the distribution appears flatter than for men 31-35 years old. However, the first and third quartiles also move rightward over the age distribution, such that the peak of the

[14] For these estimates, I use the k nearest neighbors in a centered interval around each observation, where $k = (2 * (N)^{-1/5} + .05) * N$.

wage curve is more or less in the middle of the interquartile range for men under 35, but for men over 35 the interquartile range captures more of the downward sloping portion of the distribution–to the right of the wage peak. To get a clearer sense of this, one can imagine a straight line drawn through the first and third quartile values of predicted wages: as men get older, the slope of this line gets more negative. For men 31-35, 36-40, and 41 and up, the marginal effect of an extra BMI point is to reduce wages by about .8 percent, 1 percent, and 1.8 percent, respectively.

Both of these estimates confirm the intuition conveyed by Figures 2.1 and 2.2, and our sense that our current understanding of the effect of body mass on wages may be incomplete. For women, the effect of BMI diminishes as they they grow older. For men, the effect of BMI increases as they get older. These figures also point to different age-related changes in the distribution of BMI that subtend these results. For women, the effect of BMI diminishes not simply because everyone is getting heavier, but also because as they get older, the distribution is more right-skewed. Extreme values become more prevalent and less penalized, relatively speaking. For men, on the other hand, the change in the effect of body mass on wages is largely due to the shift to the right in the BMI distribution. In addition, the univariate nonparametric models for men indicate that we should use caution in using or interpreting linear specifications of wages.

Main Regression Results

Women

The top panel of Table 2.3 shows the results of specifications of (2.1) and (2.3) for women.[15] This table shows the results from OLS and FE models using obesity indicators (with normal weight being the reference group) interacted with age dummy variables in the left panel. In the right panel are OLS and FE models using current period BMI interacted with age

[15] All of these results are for partially interacted models–that is, models that interact age only with BMI, rather than with all variables in the model. I focus on these methods because estimation of fully interacted semiparametric models is impractical in a semiparametric context. Parametric results from fully interacted models yield qualitatively similar results.

dummy variables as the variables of interest. [16]

The marginal effects of both obesity (shown in the left panel) and BMI (shown in the right panel) in both OLS and FE models have a similar pattern over the age distribution. On average, the effect of body weight increases as women get into their early 30's and declines thereafter.[17] The magnitudes of the effects are larger in the OLS than the FE models, as might be expected from a specification in which unobserved personal characteristics play a part in measuring the effect of body mass on wages.[18] Moreover, it is interesting to note that although the literature in general has found negative effects of obesity and BMI on wages for women, Norton and Han (2008) found no measurable effect of lagged BMI on wages when they used an instrumental variables approach on a sample of women who were between 21-26 years old. The results shown in Table 2.3 for women 21-25 using both obesity indicators and linear specifications of BMI are consistent with that result.

The results from these specifications suggest that on average women in their early 30's can expect to see their wages drop by about 0.5 percent for each extra BMI point (again, about 7.5 pounds for most people), and that obese women this age can expect to earn 7.7 percent less than their normal weight counterparts at this age. Women who are over 40 would see only an average 0.25 percent drop for this same increase in body mass and they may see none at all. Similarly, obese women over 40 would see their wages drop by only about 3.9 percent. Moreover, all of these estimates–those using BMI and those using obesity indicators–suggest that the difference in the effect of BMI on those who are 31-35 and those over 40 is signficantly different at $p <.10$.

Once again, this result implies an important revision of our understanding of the effect of body mass on wages for women. For the most part, researchers have assumed that BMI and obesity negatively affect wages for women in a uniform way. But these results suggest

[16]I have also estimated a FE model using a one-period lag of BMI (and age interactions with it) as the primary regressor of interest. The results of these models are essentially identical to the FE models shown here.

[17]This result is similar to that of Gregory and Ruhm (2009), who found that women 35-45 showed the greatest negative effects of BMI on wages, while those older and younger showed less effects.

[18]A Hausman test indicates that the unobserved time-invariant factors are correlated with the $X's$ and rejects a random effects specification for a fixed effects specification at $p < .001$.

not only that the effect of BMI and obesity may diminish as women age, but that as they reach their 40's there is certainly a smaller negative effect on average, and there may be no penalty to additional body mass at all.

Men

The results for the linear models for men are shown in the bottom panel of Table 2.3. All of the estimates suggest that, on average, as men get older the marginal effect of extra body mass goes from being near zero or slightly positive for those under 35 to being strong and negative for those over 40. For example, according to the results of FE models, for men who are 26-30 and 31-35, the effect of obesity on wages is relatively small and statistically insignificant: point estimates suggest that being obese increases wages by about 1.6 percent for men 26-30, while it decreases wages by about 0.7 percent for men 31-35. However, obese men 36-40 years old see a reduction in wages by 4.9 percent; obese men at least 40 years old experience a decrease in wages of about 5.7 percent. These estimates are both significant at $p < .05$.

Similarly, the models using BMI suggest increasing negative effects of BMI as men age. In FE models, the average effect of an extra BMI point is positive for men 21-25 and 26-30, but for men over age 40, the average effect of extra body mass is to reduce wages by about 0.5 percent per BMI point. (This last effect is statistically significant at $p < .05$.) OLS models using both BMI and obesity show similar patterns of effects, with significant effects emerging for men as they reach their 40's.[19]

Although these patterns in estimated effects are consistent across models, it's important to know whether any of these age group predictions are statistically different from one another. Tests show that in the FE model all the estimates of the effect of obesity for men over 25 years of age are different from the effect for 21-25 year olds at $p < .01$. Marginal effects for men 26-30 and 31-35 are different from those for men over 40 at $p < .01$ and

[19]Hausman tests confirm both the correlation of the unobserved time-invariant factors with the $X's$ and the appropriateness of the fixed effect rather than random effect specification. These tests were significant at $p < .001$.

$p < .05$, respectively. In addition, estimates for these age groups are different from those for 36-40 year olds at $p < .01$ and $p < .05$, respectively, as well. In general, all of these FE results suggest that the process determining the effect of obesity on wages changes significantly as men age.

One could say much the same thing about the OLS and FE models using BMI. The estimates of the average effects of BMI for age groups over 25 are statistically different from estimates for those 21-25 at $p < .01$. Moreover, the estimates for those 26-30 and 31-35 are statistically different from the estimates for those over 40 years of age. In the FE model, the point estimates for those 26-30 and 31-35 are significantly different than predictions for those over 40 at $p < .01$. In the OLS models, these point estimates are different from those for men over 40 at $p < .05$ and $p < .10$, respectively. In addition, in the FE model, the estimated marginal change in wages for a point change in BMI for men 26-30 and 31-35 is different from the estimated change for men 36-40 at $p < .01$ and $p < .05$, respectively.

Taken as a whole, these results strongly suggest that the previous literature understates the effect of BMI or obesity on men's wages by dint of the fact that it has not, in general, examined older workers. My results indicate that, while men in their late 20s to late 30s may not experience any detriment to wages for extra body mass on average, men over 40 experience a significant 0.5 percent decline in wages on average for each extra point of BMI. Similarly, men under 35 may experience little or no decrease in wages for being obese, but men 35-40 experience an average 4.9 percent decline in wages for being obese, and men over 40 experience a 5.6 percent decrease in wages for being obese. These estimates also indicate that differences between age group estimates are also statistically significant.

Semiparametric Results

The main advantages of conventionally used linear or piecewise linear models such as those shown above are relative efficiency and ease of interpretation. The drawbacks have to do with the imposition of functional forms. As mentioned above, there are good *a priori* reasons to suspect that linear models might not be ideal for modeling wages as a function of body mass. In general, models linear in BMI will only give an average effect, which could miss important aspects of the wage-BMI distribution, especially if low and high BMI are indicative of poor health and thus cause lower wages. Piecewise linear models (i.e. those that use weight category dummy variables) can certainly do better if one chooses appropriate cutpoints, and they can be fairly flexible if one decides to use many cutpoints, although the more cutpoints one adds, the more quickly the efficiency benefits of this strategy diminish.

I use the partial linear model here as a way of overcoming the functional form restrictions of more conventional models. While this model is less efficient in producing pointwise predictions of wages than linear and piecewise models, in the context of estimating the slope of the conditional wage function, there is less efficiency loss, as the bootstrap estimates of local slopes rely on the asymptotic characteristics of point estimates of relatively large finite samples.[20] I address the interpretability issues of the partial linear FE model sby simulating "local" slopes of the conditional wage function using bootstrapped estimates of $\hat{f}(BMI_{it})$. For women, I use the ranges 18-20, 20-25, 25-30, 30-35, and 35-40. For men, since the BMI distribution is more compact than for women and since the peak of the conditional wage function is at or near a BMI of 27, I have used ranges 19-22, 22-27, 27-32, 32-37, and 37-40. The results of these simulations for women and men are shown in the top and bottom panels of Table 2.4, respectively.

Figures 2.5 and 2.6 show results from model 2.4 for women and men, respectively. I plot this specification along with model 2.3, as well as with a model that also includes a quadratic in BMI as well as age interactions with quadratic terms. I include the quadratic to compare the partial linear models with a much more easily estimable, but also more flexible estimate with respect to BMI. The graphs also show vertical lines at the first and third quartile values, as in Figures 2.3 and 2.4.

[20]The local linear regression framework used below and the linear framework used for results in Table 2.3 also both minimize mean squared error, with the difference that the local linear framework minimizes it by the choice of bandwidth–i.e. the optimal neighborhood of observations.

Women

Visual inspection of the panels in Figure 2.5 suggests that, especially for women over 30, the parametric linear and/or quadratic models do well at approximating the distribution of conditional wages. Whatever non-linearities there were in unconditional wages for women (as shown in Figure 2.3) are largely accounted for by other controls.[21] Only estimates for the youngest two age cohorts differ markedly from the linear and quadratic models. For women 21-25 years old, however, the estimated marginal effect of BMI on wages is significant only in the BMI range from 25-30. For women 26-30, the marginal effects shown in Table 2.4 are between the OLS and FE estimates from Table 2.3. They indicate that women in this group see a decrease in wages of about 0.5 percent for each extra point of body mass. Women who would have a BMI greater than 35 at this age (who are "obese 2" by NIH standards) see a decrease in wages of about .075 percent for a similar increase in body mass. [22]

In general, the estimates of the marginal effect of BMI on wages lie somewhere in between the OLS and FE models shown in Table 2.3, but, as can be seen in the Figure 2.5 they do not differ markedly from the parametric linear or quadratic models. The results for women over 40 show an odd convexity in both the partial linear and quadratic FE models; the effect measured by the linear model in this sense is a weighted average of a strong negative effect for women in the 20-25 BMI range (-.012), negligible effects in the range from 25-30 and 30-35 (-.0045 and -.0024, respectively), and a small positive effect for women with a BMI over 35 (.0055). This odd functional shape notwithstanding, this estimate, compared to the others, reaffirms the conclusion that we draw from the above estimates. As women age, the penalty to increased body mass or obesity tends to increase until women reach their early to mid 30s and then decreases and may abate altogether for women over 40.

[21]This is consistent with results shown by Gregory and Ruhm (2009), who found that, in a comparison of conditonal wages for restricted samples in the NLSY and PSID, conditional wages were closer to being linear for women in the NLSY than the PSID.

[22]This latter result, in particular, seems driven by values of wages in the right tail of the BMI distribution. Roughly 5 percent of this sample has a BMI greater than 35; however, the marginal effect in this tail of the distribution is not statistical noise, as is suggested by the significance level of the estimate.

Men

There are two features of the estimates for men, shown in Figure 2.6, that are particularly interesting. First, these results strongly suggest that models of wages that use linear or even quadratic models miss a lot of the variation over the BMI distribution.[23] In each of the panels in the figure, the linear and the quadratic models do a good job of capturing what is happening in the middle of the BMI distribution, as can be seen by comparing the linear and/or quadratic with the semiparametric estimates in the region bounded by the first and third quartiles. However, the linear and quadratic models do less well at capturing the distribution in the bottom and top quartiles. For example, the simulated results for men 21-25 years old suggest that the marginal effect of BMI is large and positive for men in the body mass ranges from 19-22 and 22-27. The point estimates of these effects are .046 and .011, respectively, and both are significant at $p < .01$. The latter is similar to the marginal effect measured in the linear model, for which the point estimate (standard error) is .0101 (.0026). For men in body mass ranges 27-32, 32-37, and 37-40, however, the measured effects of BMI are -.0003, -.0136, and -.0167; the estimates for the last two BMI ranges are significant at $p < .01$.

Similar patterns emerge in all of the age cohorts. Below a BMI of about 27, there are positive effects of additional BMI, but above that BMI level there are large and significant penalties for additional body mass. For example, for men in the 31-35 year old age group, the estimate of the marginal effects of extra BMI for those in the BMI ranges of 19-22 and 22-27 are .0525 and .0310, respectively. However, for men whose BMI is above 27, these marginal effects of extra BMI are estimated to be -.0113, -.0041, and -.0047 for those in the BMI ranges 27-32, 32-37, and 37-40, respectively.

Second, the BMI level at which wages peak declines with age. The BMI levels at which wages are at their maximum for those aged 21-25, 26-30, 31-35, and 36-40 are 29.3, 28.1, 27.3, and 26.0, respectively. (For men over 40 years old, the peak of conditional wages is at

[23] This result is consistent with results shown by Gregory and Ruhm (2009), whose comparison of semiparametric regressions on restricted NLSY and PSID samples indicated a high degree of concavity in the NLSY79 men's sample results.

a BMI of 26.7.) At the same time, as the plotted interquartile ranges suggest, the "middle" of the distribution is trending rightward as men get older. Thus, the stronger negative effect of BMI on wages as measured by the linear model is not so much due to the penalty BMI increasing *per se*. Rather, as men get older, more of them are in the part of the wage distribution in which wages are penalized at all, partly because they are getting heavier and partly because the peak of the wage distribution is shifting to the left.

These estimates suggest that the literature has understated the wage penalty for men in two ways. First, it has not examined older workers. Second, it has relied on linear specifications that may not capture the magnitude of the wage penalty, but only give an average effect (in the linear case) or a mis-specified estimate (in the piecewise linear case). According to my estimates, wages are increasing for men with BMIs lower than about 27 and decreasing above that; as men age and have higher BMIs, the negative effects predominate. The linear FE model suggests that men who are over 40 receive about 0.5 percent less in wages for every extra BMI point. The partial linear FE model, however, suggests that for the 75 percent of men whose BMI is over 27 at this age, the effect of BMI will be to decrease wages between 0.75 and about 1.0 percent. Similarly, the partial linear model suggests that the effect of additional BMI for men 36-40 with a BMI above 27 will be between a 1 and 1.2 percent decrease, three times that predicted by the linear model.

Discussion

The literature on the effect of BMI or obesity on wages has for the most part found significant effects of body mass for women and none for men. However, most of these studies do not include older workers–those over 35 years old–and most do not examine how the effect of BMI or obesity might change with age. This study addresses this gap in the literature. In general, the findings presented here suggest that the wage penalty to body mass for women peaks when women are relatively young and diminishes after that, and that the penalty for young men is small or negligible, but large and signficant as they get beyond 40 years old.

For women, unconditional measures, linear OLS and FE regressions, as well as partial linear FE regressions point strongly to the same conclusion. As women get older, the negative effect of having extra body mass diminishes and may disappear completely. Moreover, the negative association of BMI with wages is not an obesity effect: the negative effects of extra body weight are present at all levels of BMI. Unconditional nonparametric regressions suggest that this effect is accompanied by two dynamics in the distribution of BMI: first, women are getting heavier as they get older, and, independent of that, more extreme BMI is becoming more prevalent (i.e. the BMI distribution is more right-skewed), and less penalized.

Of the unobserved mechanisms that might drive the negative effect of BMI on wages, two imply increasing penalties as women get older: decreasing marginal productivity and health care costs. My results suggest that these are likely not alone in driving the negative effects of BMI and/or obesity on wages. The effects on wages for women over 40 are less than the effects for younger women. If the health costs or marginal productivity stories were correct, we should expect to see increasing–or, at least, non-decreasing–effects as women get older. On the other hand, if the causes were appearance or consumer preferences for thinner workers, then we would expect to see what my results show, a declining effect as women age.

For men, the estimates presented here suggest a very dynamic process underlying the relationship between wage and BMI. Unconditional means and linear OLS and FE estimates point to the fact that, while extra BMI may have a positive or no effect for younger men, the effect of BMI or obesity on wages for men over 40 is unequivocally negative, large, and significant. While much of the literature has found no effect of BMI on wages for men, these results find that each extra BMI point implies a decrease in wages of about 0.5 percent for men over 40; similarly, I show that obese men over 40 earn 5.6 percent less than their normal weight counterparts, and that similar penalties are found for those in their late 30's.

My results for men also indicate that our current understanding of the effect of BMI on wages for men is almost certainly limited by linear models used heretofore in the literature.

All of the results for men suggest highly non-linear conditional wages, with positive marginal effects of BMI measured at BMI levels less than 27, and strong and negative effects for those with BMIs above 27. For the youngest men, positive effects predominate; as men age and get heavier, the negative effects become increasingly important. Simulations using semiparametric estimates of conditional wages suggest that the true effect of BMI on wages for men with BMIs above 27 may be almost three times as large as those estimated by linear models.

The results for men offer mixed evidence on the causes of the effect of body weight. On one hand, it appears that the health effects of body mass may be the cause of the wage penalty. Not only do wages begin to decline at a BMI level near the obesity cutoff, but the negative effect grows stronger as men age. However, the semiparametric estimates suggest negative effects at high levels of BMI in all age groups, which could suggest that something other than health effects–like consumer preferences or appearance–play a role for young men.

The men's results also bear on the aforementioned strand of the literature that highlights the deficiencies of BMI as a measure of body adiposity (Burkhauser and Cawley, 2008b; Wada and Tekin, 2007; Johannson et al., 2009). It is well understood that BMI does not distinguish between body fat and muscle mass, a fact that is a particular problem for understanding the effect of body fat on wages for men. Whereas we normally presume that a high BMI is evidence of high body fat, which is assumed to be detrimental to health and appearance and thus wages, men–particularly young men–who have high BMIs may be very athletic and healthy, and thus be expected to have higher–or certainly not lower–wages than their normal weight peers. But we also might expect BMI to do a better job of measuring body fat as men get older, at least if we assume that men are less physically active as they age. In this case, BMI ought to better reflect body fat as men age, and the measured effect of BMI on wages ought to better reflect the penalty to body-fatness as they age. The results shown here could suggest that there is a penalty to body-fatness for men at all ages, but BMI is not really measuring it until men reach their 40's. This would also partly explain the large positive effect of obesity and BMI on wages for young men: men 21-25 years old who are obese have 11 percent higher wages than normal weight men of the same age; men in this age group also show an increase in wages of about one percent for each BMI point. These could reflect returns to health or appearance.

These results suggest several avenues for future research. First, more research is needed into the joint relationships between wages, health expenditures, and BMI. Recent research by Bhattacharya and Bundorf (2005) suggests that the incidence of health costs explain much of the effect of BMI or obesity on wages. These estimates suggest something quite different. Future research might seek to use data such as MEPS that contains information on wages as well as health insurance, health expenditure, and health status to model these relationships. Second, these results for men suggest that one of the reasons for the increased negative association of BMI with wages for men is the better job that BMI is doing at measuring body-fatness as men age. Although it would be useful to have better measurements of body-fatness in social science data sets, as suggested by Burkhauser and Cawley (2008b), it would also be good to understand in what ways various measures fail, relative to some gold standard of measurement, such as denistometry. Both these lines of research would enhance our understanding of the effects of BMI or obesity on wages, as well as other outcomes of interest to health and labor economists.

Tables and Figures

Table 2.1: Sample Frequencies, NLSY79 Sample

Years in Sample	Women		Men	
	Persons	Person-Years	Persons	Person-Years
1	84	84	120	120
2	103	206	105	210
3	105	315	73	219
4	117	468	73	292
5	139	695	98	490
6	163	978	110	660
7	174	1,218	127	889
8	183	1,464	131	1,048
9	182	1,638	158	1,422
10	224	2,240	154	1,540
11	248	2,728	213	2,343
12	264	3,168	255	3,060
13	346	4,498	348	4,524
14	337	4,718	450	6,300
15	162	2,430	212	3,180
16	80	1,280	177	2,832
Total	2,911	28,128	2,804	29,129

Table 2.2: Weighted Sample Means in NLSY

Sample Means				
	Women		Men	
Variable	Mean	Std. Dev.	Mean	Std. Dev
Log Wage	2.529	0.566	2.795	0.592
BMI	25.708	5.687	26.646	4.345
Northeast Region	0.185	0.388	0.188	0.390
Midwest Region	0.271	0.445	0.303	0.459
West Region	0.162	0.369	0.169	0.375
SMSA	0.783	0.412	0.778	0.416
Married	0.564	0.496	0.546	0.498
Divorced/Sep.	0.179	0.383	0.127	0.333
Rural Residence	0.234	0.423	0.230	0.421
Number of Children	1.109	1.155	0.883	1.158
Hours Worked/Wk.	36.86	11.24	43.63	10.48
Job Tenure	225.80	256.36	255.89	280.75
Enrolled in School	0.072	0.258	0.055	0.228
Working Part Time	0.087	0.282	0.035	0.184
Highest Grade Completed	13.465	2.304	13.301	2.473
Total Labor Mkt. Experience (Wks.)	506.14	265.41	560.40	270.78
Age: 21-25	0.183	0.386	0.176	0.380
Age: 26-30	0.230	0.421	0.238	0.426
Age: 31-35	0.219	0.413	0.226	0.418
Age: 36-40	0.176	0.381	0.170	0.376
Age: 41 & Up	0.192	0.394	0.189	0.392
BMI @ Age 21-25	23.274	4.233	24.519	3.586
BMI @ Age 26-30	24.596	5.155	25.746	3.884
BMI @ Age 31-35	25.897	5.654	26.781	4.138
BMI @ Age 36-40	27.085	5.994	27.792	4.448
BMI @ Age 41& Up	27.865	6.047	28.560	4.523
N	28,128		29,129	

Table 2.3: Marginal Effect of BMI/Obesity on Wages at Different Ages[1]

Age Group Interaction	Models Using Obesity		Models Using BMI	
	OLS	FE	OLS	FE
	Women			
Age 21-25	-0.0663	-0.0232	-0.0065	-0.0007
	(0.0212)	(0.0242)	(0.0013)	(0.0018)
Age 26-30	-0.0872	-0.0393 ∘	-0.0071	-0.0021∘
	(0.0156)	(0.0173)	(0.0010)	(0.0014)
Age 31-35	-0.1175⋆†	-0.0768⋆†	-0.0078†	-0.0052†
	(0.0148)	(0.0160)	(0.0010)	(0.0013)
Age 36-40	-0.0788	-0.0456	-0.0062	-0.0036
	(0.0180)	(0.0173)	(0.0012)	(0.0013)
Age 41 & Up	-0.0696	-0.0394	-0.0047	-0.0024
	(0.0228)	(0.0218)	(0.0015)	(0.0015)
N	28,128			
	Men			
Age 21-25	-0.0146	0.1132	0.0025	0.0096
	(0.0194)	(0.0222)	(0.0015)	(0.0022)
Age 26-30	-0.0467	0.0159***⋆⋆⋆‡‡	-0.0027***‡	0.0023***•⋆⋆⋆‡‡
	(0.0161)	(0.0167)	(0.0014)	(0.0018)
Age 31-35	-0.0377	-0.0069***⋆⋆ ‡	-0.0032***†	-0.0002***‡‡⋆⋆
	(0.0158)	(0.0158)	(0.0014)	(0.0017)
Age 36-40	-0.0684	-0.0494***	-0.0056***	-0.0033***
	(0.0205)	(0.0187)	(0.0018)	(0.0018)
Age 41 & Up	-0.0646	-0.0566***	-0.0077***	-0.0051***
	(0.0268)	(0.0235)	(0.0020)	(0.0020)
N	29,129			

[1] Table shows the marginal effect of obesity (relative to normal weight) and BMI on log hourly earnings for women and men in the NLSY sample. Independent variables include race (in OLS models), education, total labor market experience, job tenure, hours worked per week, Census region, SMSA indicator, rural indicator, number of children, part-time indicator, school enrollment indicator, marital status, and year indicators. All models use Huber-White sandwich estimates of standard errors. Standard errors in parenthesis. † Different from Age 41 & Up at p<.10. ‡ Different from Age 41 & Up at p< .05. ‡‡ Different from Age 41& Up at p< .01. ⋆ Different from Age 36-40 at p< .10.⋆⋆ Different from Age 36-40 at p< .05 ⋆⋆⋆ Different from Age 36-40 at $p < .01$ • Different from Age 31-35 at p< .10 ∘ Different from Age 31-35 at p< .01.***Different from Age 21-25 at p< .01.

Table 2.4: Simulated Marginal Effect of BMI on Wages[1]

	Women				
	BMI Range				
	18-20	20-25	25-30	30-35	35-40
Age 21-25	-0.0007	-0.0012	-0.0025*	-0.0009	-0.0006
	(0.0063)	(0.0014)	(0.0010)	(0.0021)	(0.0016)
Age 26-30	0.0118	-0.0048**	-0.0043**	-0.0050**	-0.0075**
	(0.0088)	(0.0017)	(0.0011)	(0.0014)	(0.0023)
Age 31-35	0.0174	-0.0074**	-0.0051**	-0.0048**	-0.0018
	(0.0083)	(0.0023)	(0.0013)	(0.0009)	(0.0016)
Age 36-40	0.0002	-0.0037	-0.0052**	-0.0052**	-0.0039**
	(0.0063)	(0.0022)	(0.0008)	(0.0009)	(0.0013)
Age 41 & Up	0.0025	-0.0118**	-0.0045	-0.0024	0.0055**
	(0.0082)	(0.0035)	(0.0014)	(0.0012)	(0.0017)
	Men				
	BMI Range				
	19-22	22-27	27-32	32-37	37-40
Age 21-25	0.0460***	0.0110***	-0.0003	-0.0136***	-0.0167***
	(0.0054)	(0.0014)	(0.0014)	(0.0029)	(0.0030)
Age 26-30	0.0179**	0.0113***	-0.0073***	-0.0277***	-0.0326***
	(0.0088)	(0.0022)	(0.0019)	(0.0027)	(0.0046)
Age 31-35	0.0525***	0.0310***	-0.0113***	-0.0041***	-0.0047
	(0.0072)	(0.0022)	(0.0018)	(0.0025)	(0.0051)
Age 36-40	0.0116***	0.0171***	-0.0120***	-0.0097***	-0.0124
	(0.0075)	(0.0027)	(0.0016)	(0.0022)	(0.0059)
Age 41 & Up	-0.0017	0.0254***	-0.0073***	-0.0098***	0.0077
	(0.0196)	(0.0037)	(0.0019)	(0.0022)	(0.0039)

[1] Table shows marginal effect of BMI for specified BMI ranges for different age groups. Point estimates derived from semiparametric FE models. Independent variables include education, total labor market experience, job tenure, hours worked per week, Census region, SMSA indicator, rural indicator, number of children, part-time indicator, school enrollment indicator, marital status, and year indicators. Standard errors derived from 500 bootstrap replications; significance levels calculated using the percentile approach.

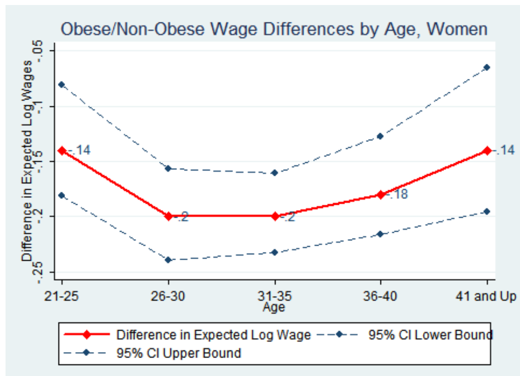


Figure 2.1: Differences in Expected Log Wages, Women

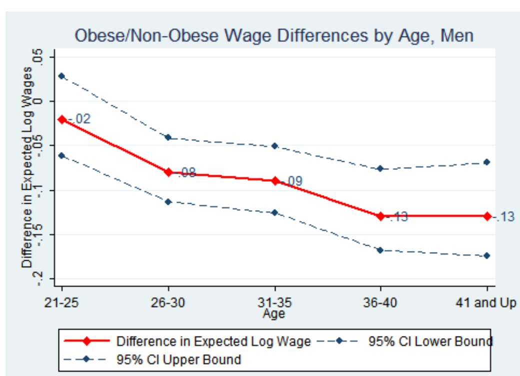


Figure 2.2: Differences in Expected Log Wages, Men

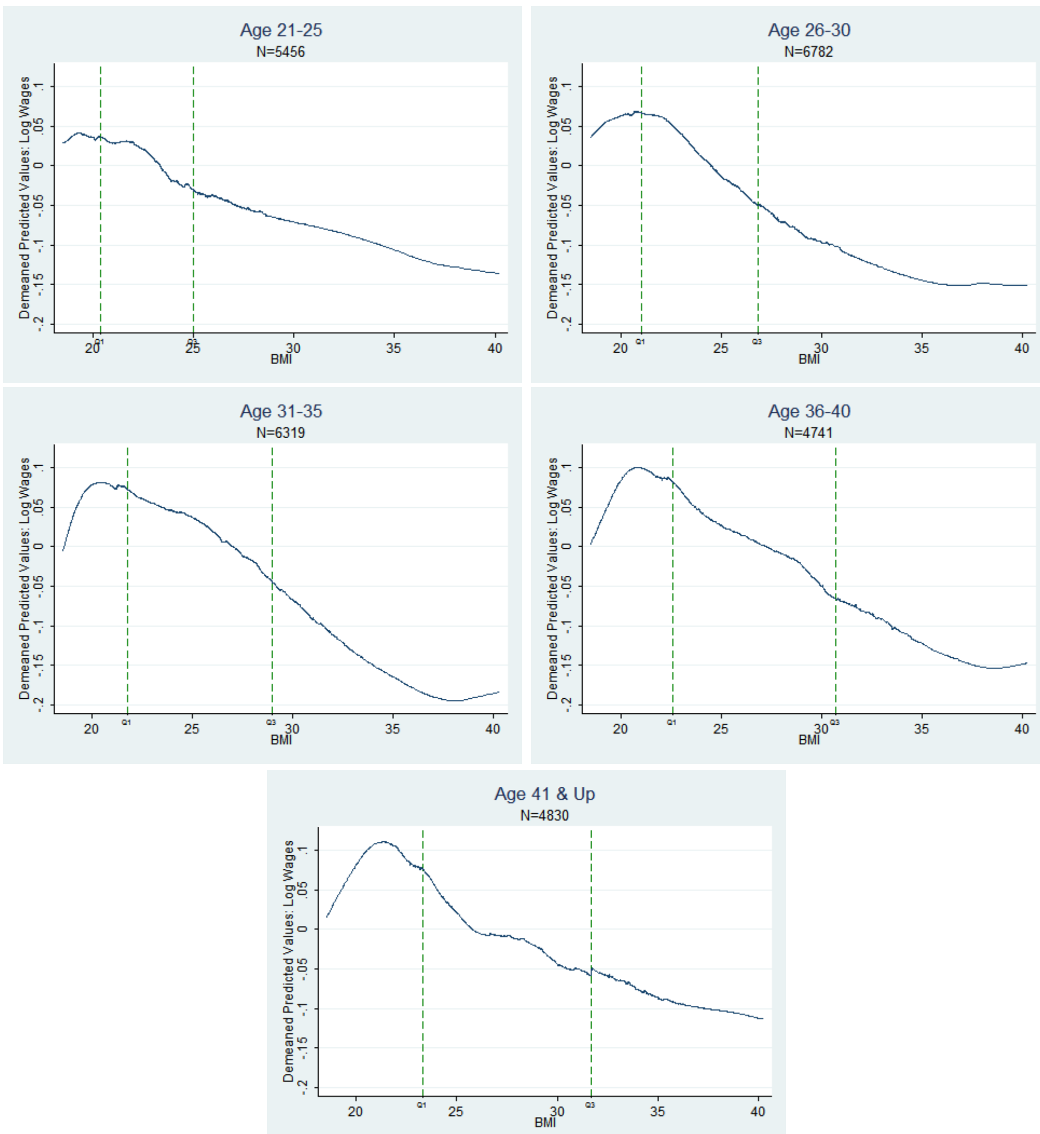


Figure 2.3: Univariate Local Linear Smoothing Estimates, Women

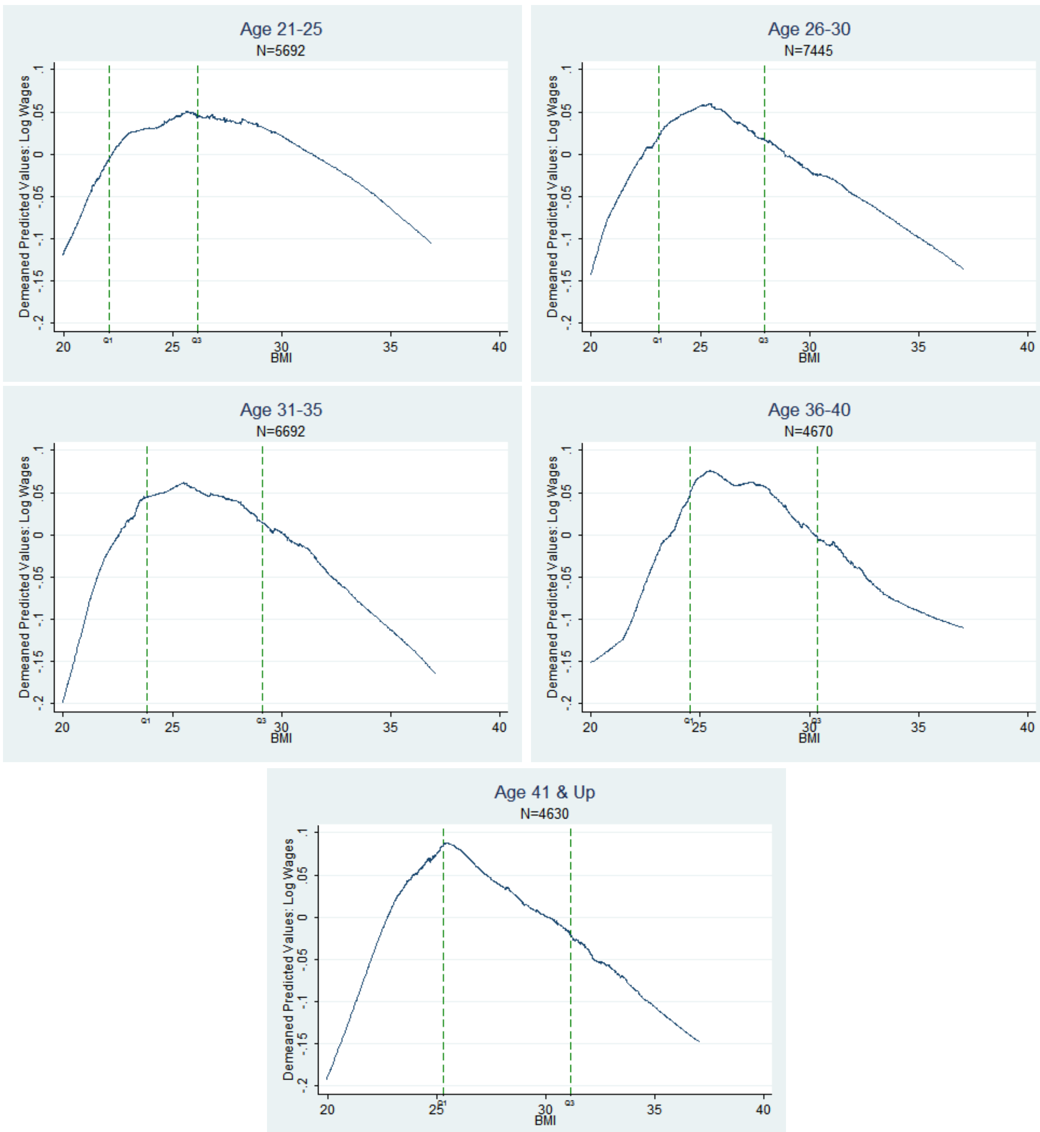


Figure 2.4: Univariate Local Linear Smoothing Estimates, Men

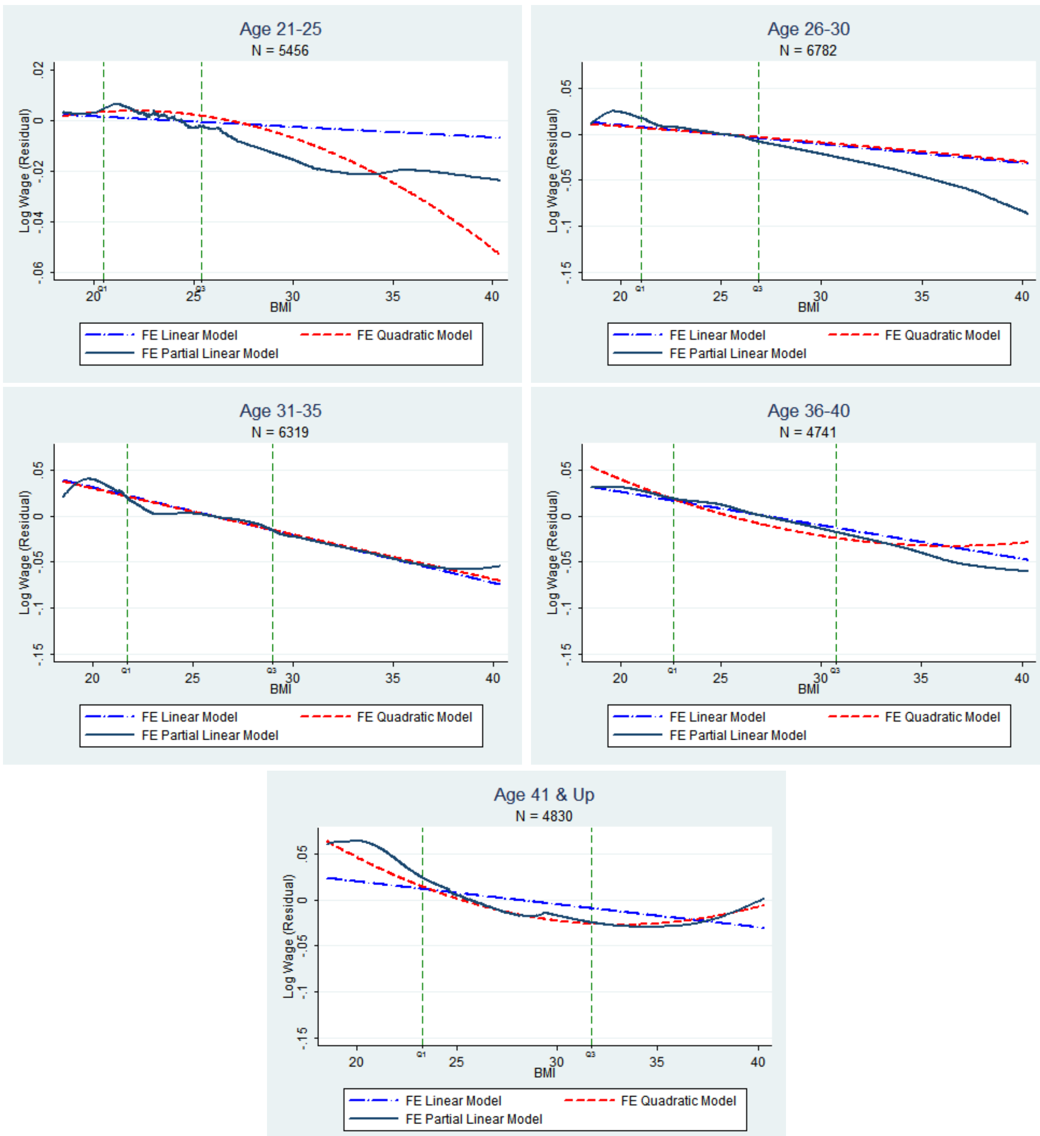


Figure 2.5: Three Models of Wages, Women

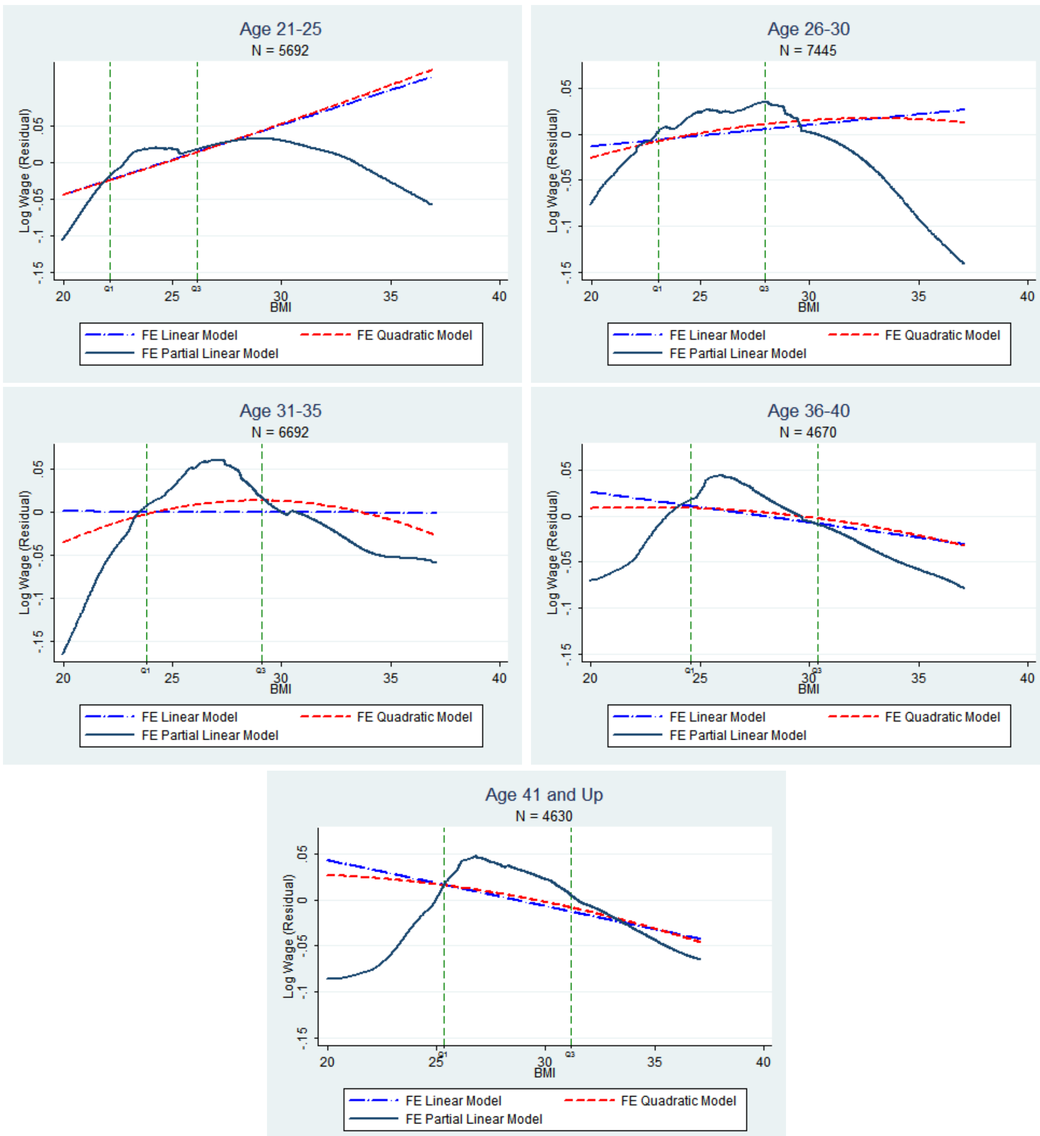


Figure 2.6: Three Models of Wages, Men

Appendix A: Sample Choice

Although the NLSY has been the workhorse data set for the examination of the effect of body weight on wages, relatively little has been said in the obesity/BMI-wage literature about its design, the choices faced by researchers using the data, and the effects of these choices on the outcomes studied. In particular, three subsamples comprise the NLSY: a representative subsample (6,185 people in 1979), supplementary subsamples (5,295 people in 1979), and a military subsample (1,206 people in 1979). The supplementary subsamples include over-samples of black, Hispanic, and economically disadvantaged white respondents. The NLSY dropped the military and economically disadvantaged white subsamples in 1985 and 1991, respectively, a fact which has led the NLSY, following the publication of a thorough review of the data by MacCurdy et al. (1998), to revise the sample weights published with the survey. Currently, for the representative and supplementary samples there are distinct sample weights: a set of weights for the representative sample only, updated for the probability of interview each wave, and a set of weights for the representative sample and all of the supplementary samples, also updated for the probability of attrition from the sample.

In this study, I employ the representative subsample for the main estimates, for the most part to retain sample consistency with the partial linear models. However, I have also estimated the main regression results with the full sample. The results for the fixed effect models from both samples are shown below in Table 2.5.

Table 2.5: Sample Choice and Marginal Effects of BMI/Obesity on Wages[1]

	Sample			
Age Group Interaction	Full	Representative	Full	Representative
	Models Using Obesity		Models Using BMI	
	Women			
Age 21-25	-0.0039	-0.0223	0.0008	-0.0007
	(0.0213)	(0.0241)	(0.0016)	(0.0018)
Age 26-30	-0.0359∘	-0.0375 ∘	-0.0016	-0.0021
	(0.0149)	(0.0173)	(0.0012)	(0.0014)
Age 31-35	-0.0706⋆‡‡	-0.0767⋆†	-0.0040 †	-0.0052 †
	(0.0143)	(0.0161)	(0.0011)	(0.0013)
Age 36-40	-0.0494	-0.0458	-0.0034	-0.0036
	(0.0154)	(0.0173)	(0.0011)	(0.0013)
Age 41 & up	-0.0209	-0.0392	-0.0014	-0.0024
	(0.0196)	(0.0218)	(0.0013)	(0.0015)
N	47,135	28,128	47,135	28,128
	Men			
Age 21-25	0.1030	0.1132	0.0089	0.0096
	(0.0196)	(0.0222)	(0.0018)	(0.0022)
Age 26-30	0.0097***⋆⋆⋆‡‡	0.0159*** ⋆⋆⋆‡‡	0.0021***•⋆⋆⋆‡‡	0.0023***•⋆⋆⋆‡‡
	(0.0143)	(0.0167)	(0.0015)	(0.0018)
Age 31-35	-0.0030***⋆⋆ ‡	-0.0069***⋆⋆ ‡	0.0006***•⋆⋆⋆‡‡	-0.0002***‡‡⋆⋆
	(0.0139)	(0.0158)	(0.0014)	(0.0017)
Age 36-40	-0.0357***	-0.0494***	-0.0020***	-0.0033***
	(0.0162)	(0.0187)	(0.0015)	(0.0018)
Age 41 & up	-0.0423***	-0.0566***	-0.0033***	-0.0051***
	(0.0211)	(0.0235)	(0.0016)	(0.0020)
N	48,865	29,129	48,865	29,129

[1] Table shows marginal effect of BMI/Obesity on wages for full and representative samples in the NLSY. Independent variables include education, total labor market experience, job tenure, hours worked per week, Census region, SMSA indicator, rural indicator, number of children, part-time indicator, school enrollment indicator, marital status, and year indicators. All models use Huber-White sandwich estimates of standard errors. Standard errors in parenthesis. † Different from Age 41 & Up at p<.10. ‡ Different from Age 41 & Up at p< .05. ‡‡ Different from Age 41& Up at p< .01. ⋆ Different from Age 36-40 at p< .10.⋆⋆ Different from Age 36-40 at p< .05 ⋆⋆⋆ Different from Age 36-40 at $p < .01$ • Different from Age 31-35 at p< .10 ∘ Different from Age 31-35 at p< .01.***Different from Age 21-25 at p< .01.

Appendix B: Missing Data

Missing wage values for persons in the sample can be evidence of selection into the labor market in the sense developed by Heckman (1979) and can be modeled as such. However, missing values for right-hand side regression variables can be evidence of a selection mechanism that is not accounted for either by longitudinal weights or labor market selection, and can therefore bias results.

There are several ways to deal with missing values of right-hand side variables. Often, economists simply delete observations for which values are not reported for any reason. This method is inefficient and can induce bias. The missing dummy variable strategy that is sometimes used (Baum and Ruhm, 2007; Cawley, 2004) will preserve efficiency, but is clearly not appropriate for continuous variables for which zero is not a valid response. There is a large and growing literature outlining methods of data imputation that make explicit assumptions about the distribution of the parameters of the data, and model the full data by Gibbs sampling, Markov Chain Monte Carlo, or related methods.[24] These methods are attractive insofar as they make explicit rather than *ad hoc* assumptions about the distribution of the complete data; moreover, although the assumption underlying these methods–that the missing data are missing at random (MAR)–cannot be tested, alternative assumptions can easily be modeled. The main drawback of this approach is that multiple imputations imply multiple sets of parameter estimates, which can be cumbersome to present and interpret.

In the main parametric models above, I have used longitudinal data to fill in as many missing values as possible–particularly those for education, hours, and age. I have discarded observations for which BMI is missing. I test the sensitivity of the results to a multiple imputation strategy as outlined in Rubin (1987) and Van Buuren et al. (1999) in Table 2.6 below. I form five complete data sets and show parameter estimates for each data set as well as for the original data used in the main body of this study.

[24] Canonical studies include Schafer (2000), Rubin (1987), and Van Buuren et al. (1999). More recent work on longitudinal data includes Daniels and Hogan (2008).

Table 2.6: Marginal Effect of BMI on Wages at Different Ages Using Imputed Data[1]

	Main Estimates	Estimates Using Imputed Data				
		Women				
Age 21-25	-0.0009	0.0002	-0.0005	-0.0010	-0.0003	-0.0005
	(0.0018)	(0.0017)	(0.0016)	(0.0016)	(0.0017)	(0.0017)
Age 26-30	-0.0020	-0.0012	-0.0021*	-0.0022*	-0.0021*	-0.0019
	(0.0014)	(0.0012)	(0.0012)	(0.0012)	(0.0012)	(0.0012)
Age 31-35	-0.0051***	-0.0038***	-0.0046***	-0.0052***	-0.0049***	-0.0045***
	(0.0013)	(0.0011)	(0.0011)	(0.0011)	(0.0011)	(0.0011)
Age 36-40	-0.0038***	-0.0026**	-0.0033***	-0.0039***	-0.0032***	-0.0034***
	(0.0013)	(0.0012)	(0.0012)	(0.0012)	(0.0012)	(0.0012)
Age 41 & Up	-0.0024	-0.0020	-0.0021	-0.0027*	-0.0024*	-0.0021
	(0.0015)	(0.0014)	(0.0014)	(0.0014)	(0.0014)	(0.0014)
N	28,128	29,057				
		Men				
Age 21-25	0.0096**	0.0091***	0.0097***	0.0089***	0.0083***	0.0084***
	(0.0022)	(0.0020)	(0.0020)	(0.0020)	(0.0020)	(0.0020)
Age 26-30	0.0023	0.0020	0.0026	0.0022	0.0010	0.0015
	(0.0018)	(0.0017)	(0.0017)	(0.0017)	(0.0017)	(0.0017)
Age 31-35	-0.0002	-0.0003	0.0006	-0.0005	-0.0016	-0.0007
	(0.0017)	(0.0016)	(0.0016)	(0.0015)	(0.0016)	(0.0016)
Age 36-40	-0.0033*	-0.0037**	-0.0028	-0.0037**	-0.0047***	-0.0037**
	(0.0018)	(0.0017)	(0.0018)	(0.0017)	(0.0017)	(0.0017)
Age 41 & Up	-0.0051**	-0.0057***	-0.0044**	-0.0051***	-0.0061***	-0.0054***
	(0.0020)	(0.0019)	(0.0019)	(0.0019)	(0.0019)	(0.0019)
N	29,129	30,015				

[1] Table shows estimated marginal effects of BMI on wages for women and men in the NLSY sample. The main estimates from the above study are shown in the first column. Other columns show results from imputed data sets. Independent variables include education, total labor market experience, job tenure, hours worked per week, Census region, SMSA indicator, rural indicator, number of children, part-time indicator, school enrollment indicator, marital status, and year indicators. All models use Huber-White sandwich estimates of standard errors. Standard errors in parenthesis. * $p < .10$, ** $p < .05$, *** $p < .01$

Appendix C: Non-Parametric Smoothing Techniques

The model that I have used for semiparametric estimates is

$$w_{it} = X_{it}\beta + [f(BMI_{it})|Age] + \mu_i + \varepsilon_{it}, \tag{2.5}$$

where Y_i is hourly wages of individual i, z_i is a vector individual characteristics and year effects, μ_i are person-level fixed effects, and $f(BMI)$ is the non-parametric function transforming BMI into wages. The method that I use is based on the "double residual" method outlined by Robinson (1988), modified for the fixed effect context as described by Li and Racine (2007). In this procedure, one first uses univariate non-parametric smoothing to purge the dependent and independent variables of the effect of BMI; second, one differences the raw variables and non-parametric predictions for each person; third, one takes the differences between the (person-level) differenced predictions and raw variables as regressors for an OLS regression. Using the estimated parameters from that equation, one then predicts $\hat{Y}_{it} = X_{it}\hat{\beta}_{\epsilon}$. With the differences between the observed and predicted values, $Y_{it} - \hat{Y}_{it} = \hat{v}_{it}$, one estimates $\hat{f}(BMI)$ using an iterative procedure. In this procedure, after getting an initial estimate of $\hat{f}(BMI)$, one updates the predictions by differencing the prediction errors over a person and then re-performing the smoothing operation. Usually, estimates converge in a few iterations.[25]

In particular, the local linear estimator that I use to estimate $\hat{f}$ can be defined as

$$\hat{r}_n(x) = \sum_{i=1}^{n} \ell_i(x) Y_i \tag{2.6}$$

where $\hat{r}_n(x)$ is the predicted value of y at a given value x, and the weights are defined by

[25]In this context, estimates converged in not more than five iterations. Li and Racine (2007) develop a method based on a likelihood profile approach, whose advantage over the local linear approach that I use here is that it will be more efficient. However, it is also quite a bit more computationally intensive, and offers no advantage in terms of consistency to the local linear approach.

the kernel function:

$$K(x) = \frac{70}{81}(1 - |x|^3)^3 I(x) \tag{2.7}$$

where

$$I = \begin{cases} 1 & \text{if } |x| \leq 0 \\ 0 & \text{otherwise.} \end{cases}$$

The choice of the kernel function–Gaussian, uniform, Epanechnikov–generally does not affect the result. The weighting function, $\ell(x)$, is defined as

$$\ell(x) = e_1^T (X_x^T W_x X_x)^{-1} X_x^T W_x$$

$$e_1 = (1, 0, 0, ...)^T,$$

where

$$X_x = \begin{bmatrix} 1 & x_1 - x \\ 1 & x_2 - x \\ 1 & x_3 - x \\ \vdots & \vdots \\ 1 & x_n - x \end{bmatrix},$$

and

$$W_x = \begin{bmatrix} w_1(x) & 0 & \cdots & 0 \\ 0 & w_2(x) & & \vdots \\ \vdots & \cdots & \ddots & \\ 0 & \cdots & \cdots & w_n(x) \end{bmatrix}.$$

Finally,

$$w_i(x) = K(\frac{x - x_i}{h}). \tag{2.8}$$

This formulation implies that the predicted value for a given value of x is the inner product of the first row of $\ell(x)$ with Y.

The choice of smoothing parameter, h, involves the tradeoff between bias and variance, as h defines the window of observations that will be used in local regression. For non-linear functions, small windows of observations give high variance and low bias, whereas large windows offer the converse. We choose the bandwidth by selecting the span, k, the fraction of the data to include in the linear estimate, to minimize mean squared error ($bias^2 + variance$) for the estimator. This implies that for each realization of x the bandwidth changes according to the distance to the observation $(k * N)/2$ observations away. In particular, I minimize the generalized cross-validation score over the range of the span as suggested by Craven and Wahba (1979). The cross-validation score is defined as

$$GCV(k) = \frac{n^{-1}\sum_{i=1}^{n}[Y_i - \hat{g}(X_i)]^2}{1 - n^{-1}tr[M_n(k)]^2} \tag{2.9}$$

where $[M_n(k)]$ is an nxn matrix with $(i,j)^{th}$ element given by $\frac{w_{ij}}{\sum_{l=1}^{n} w_{il}}$–i.e. the smoothing matrix for span k.[26]

I chose local linear regression because it relaxes the linearity assumption of OLS and minimizes both boundary bias and design bias introduced by kernel methods such as the Nadarya-Watson estimator.[27]

I use the "wild" bootstrap described by Härdle (1990). This bootstrapping procedure is appropriate for contexts in which heteroskedasticity is thought to be present, as it simulates sampling from N distributions of residuals, rather than one. At each draw of the bootstrap sample, one transforms the residuals using a two-point probability distribution, such that $\tilde{u}_{it} = [(1+\sqrt{5}/2)]\hat{u}_{it}$ with probability $p = (1+\sqrt{5})/(2\sqrt{(5)})$, and $\tilde{u}_{it} = [(1-\sqrt{5}/2)]\hat{u}_{it}$ with probability $1-p$, where $\hat{u}_{it}$ is the residual for the final iteration of the estimation procedure. One then uses these residuals in smoothing for that sample.[28]

[26]When smoothing the dependent variables, I execute generalized cross-validation at the roughly 1000 percentile points apart in the middle 95 percent of the distribution of BMI.

[27]On this point, see Wasserman (2006), 73ff., and Fan and Gijbels (1996), pp.17-18, 60ff.

[28]Although some authors (Li and Racine, 2007; Yatchew, 2003) suggest using oversmoothed predictions and undersmoothed residuals for the bootstrap sampling procedure, I do not do that here. Pilot tests suggested no difference with this procedure.

Appendix D: Main Estimation Results

Table 2.7: Main Estimation Results, Women[1]

Variable	Models Using Obesity		Models Using BMI	
	OLS	FE	OLS	FE
Black	-0.0752***		-0.0733***	
	(0.0092)		(0.0091)	
Hispanic	0.0190		0.0179	
	(0.0123)		(0.0123)	
North East	0.1390***	0.0943***	0.1389***	0.0946***
	(0.0092)	(0.0264)	(0.0092)	(0.0264)
Midwest	-0.0009	0.0038	-0.0012	0.0045
	(0.0077)	(0.0233)	(0.0077)	(0.0233)
West	0.1151***	0.1169***	0.1155***	0.1165***
	(0.0097)	(0.0260)	(0.0097)	(0.0260)
SMSA	0.1054***	0.0372***	0.1050***	0.0368***
	(0.0079)	(0.0098)	(0.0079)	(0.0098)
Married	0.0261***	0.0510***	0.0247***	0.0509***
	(0.0079)	(0.0108)	(0.0079)	(0.0109)
Rural	-0.0588***	-0.0161*	-0.0582***	-0.0161*
	(0.0085)	(0.0096)	(0.0085)	(0.0096)
Divorce/Sep	0.0295***	0.0705***	0.0275***	0.0701***
	(0.0097)	(0.0136)	(0.0097)	(0.0136)
Number of Children	-0.0170***	-0.0334***	-0.0172***	-0.0335***
	(0.0034)	(0.0047)	(0.0034)	(0.0046)
Enrolled in School	-0.0991***	-0.1492***	-0.1005***	-0.1492***
	(0.0122)	(0.0125)	(0.0122)	(0.0125)
Highest Grade Completed	0.0708***	0.0376***	0.0704***	0.0375***
	(0.0018)	(0.0051)	(0.0018)	(0.0052)
Total Labor Mkt. Experience	0.0007***	0.0007***	0.0007***	0.0007***
	(0.0000)	(0.0000)	(0.0000)	(0.0000)
Hours Worked/Wk.	0.0017***	0.0005	0.0021***	0.0005
	(0.0005)	(0.0004)	(0.0004)	(0.0004)
Age 26-30	0.0295**	0.0811***	0.0389	0.1063***
	(0.0122)	(0.0133)	(0.0399)	(0.0374)
Age 31-35	-0.0103	0.1092***	0.0159	0.2056***
	(0.0168)	(0.0202)	(0.0424)	(0.0420)
Age 36-40	-0.0606***	0.1028***	-0.0625	0.1671***
	(0.0224)	(0.0276)	(0.0487)	(0.0491)
Age 41 & Up	-0.1437***	0.0672*	-0.1822***	0.1042*
	(0.0289)	(0.0363)	(0.0572)	(0.0589)
Constant	1.1513***	1.3084***	1.2815***	1.7278***
	(0.0331)	(0.0868)	(0.0436)	(0.0796)
N	28,128			

[1] Table shows regression coefficients for regressors other than BMI- or obesity-age interations. Year indicators not shown. *p< .10, **p< .05, ***p< .01.

Table 2.8: Main Estimation Results, Men[1]

Variable	Models Using Obesity		Models Using BMI	
	OLS	FE	OLS	FE
Black	-0.1668***		-0.1663***	
	(0.0103)		(0.0103)	
Hispanic	-0.0818***		-0.0804***	
	(0.0125)		(0.0125)	
North East	0.1255***	0.0584*	0.1254***	0.0593*
	(0.0092)	(0.0303)	(0.0092)	(0.0303)
Midwest	0.0055	-0.0590***	0.0045	-0.0598***
	(0.0078)	(0.0207)	(0.0079)	(0.0207)
West	0.0974***	0.0505*	0.0971***	0.0492*
	(0.0101)	(0.0274)	(0.0101)	(0.0274)
SMSA	0.1292***	0.0272***	0.1290***	0.0269***
	(0.0084)	(0.0098)	(0.0084)	(0.0098)
Married	0.1611***	0.0923***	0.1643***	0.0929***
	(0.0087)	(0.0099)	(0.0087)	(0.0099)
Rural	-0.0575***	-0.0363***	-0.0569***	-0.0360***
	(0.0086)	(0.0091)	(0.0087)	(0.0091)
Divorce/Sep	0.0656***	0.0372***	0.0676***	0.0374***
	(0.0110)	(0.0136)	(0.0110)	(0.0136)
Number of Children	0.0295***	0.0176***	0.0289***	0.0175***
	(0.0038)	(0.0041)	(0.0038)	(0.0041)
Enrolled in School	-0.2205***	-0.2673***	-0.2233***	-0.2676***
	(0.0154)	(0.0155)	(0.0150)	(0.0155)
Highest Grade Completed	0.0757***	0.0465***	0.0757***	0.0464***
	(0.0016)	(0.0063)	(0.0016)	(0.0063)
Total Labor Mkt. Experience	0.0005***	0.0010***	0.0005***	0.0010***
	(0.0000)	(0.0001)	(0.0000)	(0.0001)
Hours Worked/Wk.	0.0002	-0.0030***	0.0005	-0.0030***
	(0.0004)	(0.0005)	(0.0004)	(0.0005)
Age 26-30	0.0641***	0.0867***	0.1919***	0.2513***
	(0.0140)	(0.0137)	(0.0540)	(0.0467)
Age 31-35	0.0584***	0.1018***	0.2072***	0.3276***
	(0.0182)	(0.0200)	(0.0555)	(0.0499)
Age 36-40	0.0562**	0.1045***	0.2639***	0.4015***
	(0.0250)	(0.0280)	(0.0649)	(0.0599)
Age 41 & Up	0.0315	0.0700*	0.2990***	0.4169***
	(0.0339)	(0.0366)	(0.0750)	(0.0683)
Constant	1.0573***	1.9198***	1.2710***	1.7054***
	(0.0429)	(0.0861)	(0.0494)	(0.0982)
N	29,129			

[1] Table shows regression coefficients for regressors other than BMI- or obesity-age interations. Year indicators not shown. *p< .10, **p< .05, ***p< .01.

CHAPTER III

IS IT GETTING EASIER TO BE OBESE?

Abstract

Improvements in treatments for co-morbidities of obesity–high cholesterol, diabetes, sleep apnea and heart disease–over the last 25 years have made obesity less burdensome. In this study I use data from the NHIS to examine whether such improvements have been borne out in obese persons' self-reports of health. My results suggest that between 1982 and 1996 obese women enjoyed significant gains in health relative to their normal weight counterparts. However, these gains do not appear to be due to improvement in treatment for co-morbidities of obesity; rather, income and especially education explain a large share of these health trends. For men, there seems to be little in the way of trends during these years of the survey. Results from the later years of the NHIS survey (1997-2006) suggest very little in the way of trends in self-reported health for obese men or women, but they suggest very large and significant improvements in health for obese women with coronary heart disease and obese male diabetics. All of these results should be interpreted with caution, as evidence of reporting anomalies in health appear to be present.

Introduction

Obesity has become a major focus of health policy in the United States, as its prevalence has increased dramatically in the last 30 years. The CDC estimates that currently 32 percent of men and 35 percent of women are obese, more than double the prevalence of the 1970's (Flegal et al., 2010). Much of the public policy concern about this increased prevalence is due to the fact that adult obesity is known to be correlated with higher mortality, diabetes, gallbladder disease, cardiovascular disease, hypertension, sleep apnea, osteoarthritis, and some cancers (Must et al., 1999; Flegal et al., 2005).

In the last decade there has been an outpouring of research examining the causes of the upward secular trend in obesity prevalence. Much of this literature is devoted to showing that the sustained increase in obesity is driven by the falling relative costs of food and the increasing costs of physical activity (Lakdawalla and Phillipson, 2002, 2007; Chou et al., 2004; Cutler et al., 2003). A good deal of the policy discussion around obesity addresses these two factors through recommendations for diet and exercise (Mello et al., 2006; Gregg and Guralnik, 2007). This is particularly true insofar as health policy with respect to obesity has been targeted at children.

One relatively unexamined aspect of the continued rise in obesity prevalence is the possible reduction in the health cost (understood in the broadest sense) of *being*–rather than becoming–obese. While we know that obese persons are more likely to suffer from coronary heart disease, hypertension, osteoarthritis, sleep apnea and some kinds of cancer than their non-obese counterparts, in general we do not know whether it has gotten easier to bear these conditions over time. One recent study has shown that the prevalence of high cholesterol has decreased significantly more for obese than normal weight persons between the first National Health Examination Survey (NHES) in 1960 and more recent National Health and Nutrition Examination Survey (NHANES) in 1999-2000 (Gregg et al., 2005); another found evidence the risk of cardiovascular disease related mortality had decreased for the obese in the cohort in NHANES III relative to the one in NHANES I (Flegal et al., 2007). These findings suggest that, in more general terms, it's possible that the health effects of being obese may have become less severe over the last 25 years, due perhaps to advances in the treatment of co-morbidities such as diabetes, hypertension, arthritis and coronary heart disease. To the degree that this is true, then the continued upward trends in obesity in part represent a rational response to the falling cost of remaining obese.

For this study, I use data on self-reported health status from the National Health Interview Survey (NHIS) to examine to whether the perceived health burden of being obese has changed since the early 1980's. Although it is an imperfect measure of underlying disease burden, self-assessed health is arguably as important as more objective measures of health for understanding behavior and making health policy. Most importantly, self-assessments of health affect willingness to adhere to treatment regimes and undertake lifestyle changes that affect the long-term costs of chronic illnesses associated with excess bodyweight. Moreover, such self-assessments also affect labor market, fertility, and education decisions, which have important cost and productivity implications.

In examining trends in self-reported health, I make use of data on self-reported "objective" health conditions in the NHIS. I use health condition measures–indicators for coronary heart disease, heart attack, hypertension, diabetes, and arthritis–as mediating variables in explaining trends in health for obese persons. The intuition behind their use is that, if treatments for these conditions has improved, and if obese persons are more likely to suffer from these conditions, then they ought to benefit disproportionately from improvements in treatment. Moreover, improvements in these treatments ought to explain any trends in global health for obese persons–which I expect to show improvement–over these years.

My results suggest that between 1982 and 1996 obese women enjoyed significant gains in health relative to their normal weight counterparts. However, these gains do not appear to be due to improvement in treatment for co-morbidities of obesity; rather, income and especially education explain a large share of these health trends. For men, there seems to be little in the way of trends during these years of the survey. Results from the later years of the NHIS survey (1997-2006) suggest very little in the way of trends in self-reported health for obese men or women, but they suggest very large and significant improvements in health for obese women with coronary heart disease and obese male diabetics. All of these results should be interpreted with caution, as significant anomalies in health trends and interactions appear to bes present.

Data

National Health Interview Survey: Design

I use the 1982-2006 waves of the National Health Interview Survey (NHIS). The NHIS is an annual national probability sample survey of the non-institutionalized population of the United States. The survey collects information on demographics, health conditions, and, in select years, health behaviors and health insurance for all sample respondents. Information on self-assessments of health, height, and weight for adult sample respondents is available in all years of the survey.[1]

The NHIS has a stratfied, multistage design for which minor changes between 1982 and 2006 took effect in 1985, 1995, and 2005. A major redesign of the survey took place between the 1996 and 1997 survey waves. There are three important design differences because of which I treat the years 1982-1996 and 1997-2006 separately in this analysis.[2]

First, until 1996, the NHIS sampled all persons over 18 years old in each household either through direct interview or through the report of another adult in the household; after 1996, however, NHIS collected information on a single, randomly chosen adult in the household. Thus sample sizes in the post-1996 waves are less than half than those for the earlier years; the variance estimates are accordingly larger. Additionally, the change in sample design also creates differences in the composition of the sample that will affect the measurement of trends in self-reported health. Such differences are illustrated in Table 3.1, which shows the means of educational attainment, income, and health variables across the sample re-design years 1996 and 1997 along with the mean yearly growth in those variables for the years up to and including 1996 and for the years after 1996. The top panel shows changes for women, the bottom for men. For each of these variables, the relevant comparison is of

[1] Before 1982, self-assessed health was recorded on a four-point scale. Beginning in 1982, it was recorded on a five-point scale.

[2] Variables to identify sampling probability (weights) and primary sampling units (PSU's) are available in all years. Stratification information has been available in the public use data since 1987. Recently, the Integrated Health Interview Series (`http://www.ihis.us/ihis/`) has published survey stratification data for the NHIS from 1982 to 1987. I use that data here.

the year-on-year growth from 1996 to 1997 (shown in column 3) with the average growth rates of these variables before and after the re-design (columns 4 and 5).[3] For each of the variables shown here, the changes over the single year 1996-1997 are several times larger than the average rates of growth (or decay) either before or after the survey design change. For example, the proportion of women for whom the highest educational attainment is high school completion drops 7 percentage points between 1996 and 1997, a drop of rougly 19%, while the mean drop before and after the survey design change is 1.3% percent per year. At the same time, the proportion of women who claim some college education as their highest educational attainment increases by 5.5 percentage points (24%) between 1996 and 1997, substantially faster than the trend before or after the design change. Average income for women, measured as a percentile of the yearly income distribution, grew more rapidly across 1996 and 1997 than previous or subsequent years by at least an order of magnitude. And, as we might guess from the aforementioned changes, the likelihood of reporting very good health increased roughly 10 times faster in these years than in the years before and after the design change. Similar remarks could be made about comparisons for men. In short, the sample after the re-design is more educated, wealthy, and healthy than the sample before the redesign in a way that seems inconsistent with average growth before and after the design change. Although controls for these variables will mitigate these effects in multivariate models, in terms of measuring raw trends in self-reported health, these changes will bias estimates.

Second, although the question about self-reported health status is itself identical across years, the survey context for the question is fairly different between waves up to and after 1996. In the years between 1982 and 1996, the question on global health comes after a series of questions about injuries in the previous two weeks and visits to the doctor in the past 12 months. After 1996, the question about overall health is framed by questions about conditions causing limitation of activity. These differences may affect the assumed referent of the question about overall health.

[3] Average growth rates calculated using the geometric mean of the year-on-year survey means of respective variables.

The third change has to do with collection of information on health conditions, which, as mentioned above, I use as mediating variables to explain the correlation between obesity and good health. Prior to 1997, each household was assigned 1 of 6 condition lists–skin/musculoskeletal, impairments, digestive, miscellaneous, circulatory, or respiratory; each member of the household was asked whether they had any of the conditions on that list. Reports of conditions in addition to those on the list assigned to the household are recorded in reference to questions about limitations of activities, restriction of activities, doctor visits in the previous 2 weeks, and hospital episodes.[4] After 1996, the household adult is asked whether their doctor or other health professional had ever told them that they had any of a short list of conditions that includes hypertension, coronary heart disease, angina, heart attack, emphysema, asthma, ulcer, cancer and diabetes. Questions about limitation of activities are asked of everyone in the survey, and, in the context of those questions, other conditions that might cause limitations can be reported, but the list of these conditions is not open-ended, as in the years before 1997. For most of the conditions that I concentrate on here–hypertension, coronary heart disease, heart attack, and diabetes–the waves until 1996 include self-reports conditional on restriction, limitation, hospital episode or doctor visit, while the waves after 1996 do not. For one condition–arthritis–the self-reports before 1996 may or may not be conditional on restriction, limitation, hospital episode or doctor visit while those after 1996 are all conditional on functional limitation. These self-reported conditions before and after the survey design change are thus not comparable. Using them as controls in a single model would make interpretation of marginal effects impossible.

Dependent and Independent Variables

Self-Reported Health

I concentrate on self-assessments of health as the primary outcome of interest. For all years of the survey, this assesment is a response to the question, “Would you say that your health

[4]In the NHIS, the reference period for a restriction on activity is the previous 2 weeks: for example, influenza or kidney stones might cause a restriction on activity. A limitation on activity, however, refers to the functional state of respondent. Arthritis, for example, might limit one’s ability to work or keep house.

in general is excellent, very good, good, fair or poor?" Although it is not an objective measure of health, research has shown that self-reported health is strongly correlated with other health outcomes. For example, it has been shown to be predictive of mortality for both the elderly and the non-elderly (See, for example Kaplan et al., 1993; McGee et al., 1999), changes in functional ability (Idler and Kasl, 1995), likelihood of illness absence from work (Marmot et al., 1995), use of physician services (Miilunpalo et al., 1997), and general physical functioning (Ford et al., 2001). In addition, self-reported health is also known to be correlated with factors such as neighborhood quality, social capital (Kawachi et al., 1989), perceived health relative to a peer group and health behaviors (Krause and Jay, 1994).

Body Mass Index and Obesity

I construct body mass index (BMI) by dividing weight in kilograms by height in meters squared. I calculate BMI for this sample by using self-reported measures of height and weight. As body weight and height are known to be reported with error, I adjust these using the regression correction suggested by Lee and Sepanski (1995) and commonly employed in the literature (Cawley, 2004; Chou et al., 2004; Lakdawalla and Phillipson, 2007). Specifically, using data from the National Health and Nutrition Examination Survey (NHANES) III (1986-94), NHANES 1999, NHANES 2001, and NHANES 2003, I regress measured height (weight) on self-reported height (weight), its square and its cube. The coefficients, for models stratified by gender and race, are used to predict actual BMI (in the NHIS) as a function of self-reported BMI.[5] I use this corrected measure to form the classifications that I use throughout the analysis: underweight ($BMI < 18.5$), normal weight ($18.5 \leq BMI < 25$), overweight ($25 \leq BMI < 30$) and obese ($BMI \geq 30$).

Health Conditions

I use self-reports of five conditions known to be strongly correlated with obesity–hypertension, diabetes, heart attack, coronary heart disease, and arthritis–as controls for underlying

[5]I use multiple waves of NHANES so that I can restrict the age range of the prediction samples to those relevant to this study, namely those 35-60.

health, which, with the appropriate interactions, I expect will explain trends in self-reported general health for the obese relative to those of normal stature; I discuss this intuition below. As has been mentioned, for the early (1982-1996) and later (1997-2006) waves of the survey, these self-reports are responses to different questions. In the earlier waves of the survey, a report of one of these conditions could be a response to a general question asked about a list of conditions randomly assigned to the household.[6] It could also be a response to a question about the condition responsible for a limitation of activity, restriction of activity, a doctor visit in the previous 2 weeks, or a hospital stay. By contrast, in the later waves of the survey, the reports for all conditions except arthritis are a response to the general question "Have you ever been told by a health professional that you had (hypertension, diabetes, coronary heart disease, heart attack)?" A report of arthritis indicates that the respondent has answered affirmatively to the question, "Are you limited in any way in any activities by physical, mental, or emotional problems?" The respondent would then indicate arthritis as the cause of this limitation. Survey questions to which self-reports of conditions can be responses are shown in Table 3.2.

While self-reports such as the ones I use here have been used to address measurement error in studies that assess the income gradient in self-reported health, these condition measures are also known to be subject to error. There are two kinds: false negatives, which occur whenever someone who actually does have a condition (as might be verified by a health professional'ss examination) fails to report it, and false positives, which occur when someone claims to have a condition that they do not have. These errors concern me to the degree that they might be trending with self-reported health and correlated with BMI status, since such correlations would affect the explanatory power of these conditions in a regression setting. Additionally, I'm concerned about the types and causes of errors, since they will affect interpretations of models that include conditions and trends in conditions.

[6] Arthritis belongs to the skin/musculoskeletal, diabetes to the miscellaneous, and coronary heart disease, heart attack, and hypertension to the circulatory list.The question used to elicit the presence of diabetes or arthritis for those assigned the respective condition lists is "During the past 12 months, did you have (arthritis, diabetes)?" For heart attack, hypertension, and coronary heart disease, the respondent assigned to answer questions about the circulatory system is asked "Have you ever had (heart attack, hypertension, coronary heart disease)?"

False negatives occur for any number of reasons, but two are particularly relevant here. First, false negatives are likely when conditions are largely asymptomatic: hypertension is a good example of such a condition; second, they may occur if someone who is receiving treatment for a condition considers the abatement in measured effects synonymous with the absence of the condition. For example, if one were being treated for hypertension with ACE inhibitors or ARB's, one might fail to report hypertension in the survey because one conflates having normal blood pressure with medication with the absence of hypertension itself. This second kind of false negative is of particular concern because it might be trending if more hypertensives are getting treated for their condition. Although without validation data it is impossible to know for sure if errors are trending over time, we can examine trends in reported hypertension over time to see if they meet expectations given trends in objective measures from other studies.

Figure 3.1 shows the prevalence of self-reports of hypertension and the probability of reporting very good health for obese women (top panel) and men (bottom panel).[7] I plot the reports of hypertension along with very good health to show the degree to which the condition reports (and possible errors) might be trending with health for obese persons. In the early years of the survey (pre-1997), reports of hypertension and very good health trend strongly in opposite directions, especially for women. For the later years (post-1996) there seems to be less of a trend in very good health, but for both men and women a substantial increase in reported hypertension. These trends in self-reported hypertension coincide with results from studies that have examined trends in measured hypertension in periods roughly corresponding to the survey waves in NHIS. One study found that measured hypertension fell significantly between NHANES II (1976-80) and NHANES III (Phase 1: 1988-1991) (Burt et al., 1995), while another more recent study found that measured hypertension increased between NHANES III (Phases 1 and 2: 1988-1994) and NHANES 1999-2004 (Cutler et al., 2008). Both of these studies found that treatment and awareness

[7] The difference in the prevalences shown in the pre- and post-redesign waves is in part due to the question asked of the respondent: in the early years, the question refers to the previous 12 months, while in the later years it refers to the respondent's entire lifetime.

of hypertension had increased over the respective NHANES waves. If false negatives of the second variety were a serious problem in NHIS, we would expect to see self-reports trending more negatively than actual hypertension. In the early waves of the survey, we can not really distinguish trends in errors from actual trends, since both imply downward trends in self-reports as are shown in the early waves in Figure 3.1. In the latter waves of the survey, however, if false negatives are trending, we would expect to see weak or no upward trend or a downward trend in hypertension self-reports. As the figure shows, there is a fairly strong upward trend in self-reports of hypertension in these years. This is encouraging, as it indicates that, while false negatives are certainly present in the data, they may not be trending in such a way as to bias estimates of trends in self-reported health.[8]

False positives might appear for a number of reasons. People might be lying to justify their labor market status to themselves or others; they may have misunderstood their doctor; they might misunderstand the survey question, or they might self-diagnose their condition.[9] These are a source of concern to the degree that errors might be trending over time, which only seems likely if they are due to the first reason mentioned above. If, for example, women are feeling more pressure to be in the labor market over the early years in the survey and need justification to remain out of it, false positives could be trending. We might see something similar if persons are feeling the need to justify retiring early. However, controls for labor market participation in multivariate models discussed below suggest that false positives are not a primary concern.[10]

[8] In addition, one study that has examined the effect of false negatives on the income gradient in hypertension found that self-reports showed no gradient, but that objective measures did (Johnston et al., 2009). Probit models that use NHIS self-reports of hypertension as the dependent variable are strongly income graded.

[9] Two recent validation studies using data from Britain and Canada have found that the rate of false positives was less than 5 percent (Johnston et al., 2009; Baker et al., 2004), although Baker et al. (2004) found that false positives were responsible for most of the bias in coefficient estimates when self-reports are used as an explanatory variable.

[10] Models that include labor market status show no differences with the models that I show below, which do not include it.

Other Variables

As mentioned above, self-reported health has been shown to be affected by neighborhood and social characteristics. I use geographic proxies for peer health and peer attitudes toward body weight ("stigma") in selected models below. I construct the first by using the average level of health in the primary sampling unit (PSU) to which a person belongs.[11] I construct the second by using the average BMI in the PSU to which a person belongs.

Other independent variables include race, education, marital status, Census region, urban residence and income. This last I code as a percentile to accomodate its treatment as a categorical variable in the survey. In addition, for all waves after 1996, I use information on health behaviors available in the public use files. I use indicators of whether or not someone smokes and a variable indicating the number of times per week someone engages in moderate exercise for at least 20 minutes as measures of health behavior.

Means of all analysis variables for women and men are shown in the two left and two right panels of Table 3.3, respectively. In each set of panels, means are shown for the early (1982-1996) and later (1997-2006) waves of the survey, respectively.

Sample Construction

I include only persons who responded for themselves (before 1997), had valid height and weight measurements, and were between the ages of 35 and 60 in the examination sample.[12] I restrict the age group on the high end to limit mortality selection effects for the elderly. On the low end, I restrict the sample to those more likely to experience the health effects of obesity's comorbidities, which are felt by those nearing 40 years old (See Finkelstein et al., 2007). For the 1982-1996 waves, there are 183,947 (107,367) women (men) in the sample; for the 1997-2006 waves, there are 68,931 (58,389) women (men).

[11] In the NHIS, a PSU is usually one or two contiguous counties.

[12] I drop persons who report BMI below 13 or above 60 as these values appear to be the result of reporting error.

Obesity and Self-Reported Health Conditions

This study examines trends in self-reported health, focusing particular attention on the comparison between trends for persons who are obese and those who are of normal stature. I expect that trends in self-reported health for obese persons (relative to those of normal stature) to be positive, reflecting improved treatment for health conditions known to be particularly correlated with obesity: coronary heart disease, diabetes, hypertension, heart attack, and arthritis. Although the correlation between obesity and these conditions is well established in the medical literature, it is important to establish that this correlation holds for the data used in this study, which are comprised of self-reports.

Tables 3.4 and 3.5 show the simple probabilities of each of the 5 conditions examined conditional on BMI status. Broadly speaking, these tables show that the self-reports of health conditions behave as we might expect objective measures of these conditions to behave. The probability of each of these conditions is greater for obese than normal weight persons. Interestingly, for all of these conditions except hypertension for women in the early years, each of the probabilities for underweight are higher than for normal weight persons, reflecting the fact that underweight is also associated with relative detriment to health. It is also worth noting that for men in the later years of the survey, the probabilities of heart attack, coronary heart disease, and arthritis are higher among underweight than obese men. This could suggest that health status (i.e. surviving a heart attack) is driving BMI for underweight men, although these results should be interepreted with caution because the sample size for underweight men in the later years of the survey is very small (93 persons).

Difference-in-Difference in Self-Reported Health

Tables 3.6 and 3.7 represent a first pass at understanding whether it has indeed gotten easier to be obese over the years of the NHIS survey examined here. These tables show simple probabilities of reporting excellent or very good and fair or poor health in the beginning and ending years of the two series of waves: 1982 and 1996 in Table 3.6, and 1997 and 2006 in Table 3.7. The table shows the differences between normal weight ($18.5 \leq BMI < 25$) and obese ($BMI \geq 30$) persons in the beginning and ending years of these respective waves in columns 3 and 6, and then the difference-in-differences (DD) between the beginning and ending years in column 7. If the health burden associated with obesity declines over time, we expect to see the difference-in-difference estimate for Excellent or Very Good Health be negative, reflecting the fact that the normal weight-obese difference on this measure will be positive, but that this difference will be smaller in the last than the first year. Conversely, if health is improving for obese relative to normal weight persons, the DD estimate for fair or poor health will be positive, reflecting that the normal weight-obese difference will be negative, and, if relative health has converged, the last year estimate will be smaller in absolute value than the first year estimate.[13]

The results shown in Table 3.6 suggest that health has improved for obese women relative to their normal weight counterparts in the earlier waves of the NHIS survey. As we expect if health for obese persons has improved, the DD estimates for very good or excellent and fair or poor health are negative and positive, respectively, and relatively large, although the former is not statistically significant. For men in these years, obese men appear to have lost ground relative to their normal weight peers at least at the top of the health distribution, as the DD estimate for very good or excellent is large and positive (although not significant). The DD estimate for fair or poor is small and positive, offering some weak evidence that health may have improved at the lower end of the health distribution.

The results shown in Table 3.7 suggest that there has been little appreciable trend in health for obese relative to normal weight persons in the later years of the survey. Although the DD estimates for women are both of the appropriate sign, they are small and statistically insignficant. The DD estimates for men, on the other hand, suggest a slight deterioration in health for obese relative to normal weight men in the later years of the survey.

[13] I leave out good health here because it can be inferred from these two measures: fair or poor health = 1 - (excellent, very good or good health). Tables show the estimate for excellent or very good health.

Conceptual Framework and Empirical Strategy

The primary hypothesis of this study is that trends in self-reported health ought to show more improvement for obese than normal weight persons. The main reason to expect this is that changes in the technological environment have made conditions highly correlated with obesity (as well as excess weight itself) easier to treat and manage. Obese persons ought to have benefitted disproportionately from these changes, and their self-reports of general health should reflect this fact. In the years in examined in this study, advances in the treatment for obesity's comorbidities were numerous. For example, between 1988 and 1992, the first of the cholesterol lowering agents known as statins (lovastatin, pravastatin, simvastatin) were introduced to the US market: these drugs were more efficient and less toxic in managing high cholesterol than the previous generation of drugs (Endo, 1992). There have been other pharmaceutical developments that have made hypertension, congestive heart failure (ACE inhibitors, ARBs) and diabetes (orlistat, Metformin) much easier to manage (Torgerson et al., 2004; Diabetes Prevention Program Research Group, 2002). The development of portable machines to treat sleep apnea has improved outcomes with respect to sleep and blood pressure over the last 20 years (Becker et al., 2003). In addition, over the last ten years, there has been increased use of bariatric surgery to eliminate excess weight, and research into the effects of this surgery has shown that it also produces lasting results in addressing type 2 diabetes (Folli et al., 2007). While this particular procedure reduces obesity–and not its attendant health costs–it could also have the effect of reducing the expected costs of obesity, insofar as it can be relied upon to eliminate excess weight and its risks.

There are other reasons to suspect that the effect of obesity on health would diminish over time. First, the social environment has become more accomodating to the obese over time in part because of its increasing prevalence. Although such changes aren't strictly measurable, we can expect them to have an effect on health in the sense of global well-being. Second, it is well known that education is trending strongly during this period of time, especially for women. If the increases in education are particularly concentrated among the obese, this difference in trends could have a strong effect on differential trends in health, as education (as a marker of socioeconomic status) is strongly related to health outcomes.

Some caution is required in interpreting the measured relationships between global health, body mass index, self-reports of health conditions, and other covariates such as education. One concern is that self-reports of objective conditions might be endogenous to obesity and yet both of these are explanatory variables in this analysis. As I explain below, I control for variables thought to be endogenous to body mass in separate specifications and examine the changes in the marginal effects of the main right-hand side variable (an obesity-trend interaction) with and without these variables. In addition, I examine a model in which all of these variables potentially endogenous to body mass are present. While this will not address whether these variables are endogenous to obesity, it does net out their potential confounding effect.

A second serious concern is the endogeneity of obesity to self-reported health. Self-reported health contains subjective information that may be reasonably thought to be correlated with obesity. In a context in which panel data or instruments are available, such concerns can be modeled using fixed effects methods or in a structural equations framework consistent with a human capital model of health production (Grossman, 2000; Yang et al., 2005; Rosenzweig and Shultz, 1983). In the current context, neither panel data nor instruments are available, and so this concern is not directly addressable. However, since I am modeling trends in self-reports of health, so long as the distribution of unobservables across the BMI distribution is constant across time, the point estimates of the marginal effects of obesity will be unbiased.[14]

In an empirical framework, I model self reported health as a function of changes in the severity of underlying health conditions, social attitudes (stigma), neighborhood health, education, income, demographics such as age and race, BMI status (underweight, overweight,

[14]Standard errors will be too small, although by an unknown fraction of the measured variance.

obese), and health behaviors (in later years). I estimate binary logit models for which the the dependent variable is an indicator of reporting good, very good, or excellent health, and the outcome of interest is the marginal effect of the trend-obesity interactions on the right hand side of the regression model.[15] In particular, I estimate

$$H_i = \Lambda(x_i\beta), \tag{3.1}$$

where H_i is a binary indicator so described, Λ is the logistic cdf and $x_i\beta$ includes covariates as specified below.

I use a regression procedure in which I add covariates to a very simple specification to check the robustness of the marginal effects of trend-obesity interactions to the inclusion of potentially endogenous and/or mediating variables. I estimate 7 models, beginning with a regression using only a linear time trend, indicators for underweight, overweight and obesity, and interactions between the linear trend and each of the indicators: this is model 1 shown in Table 3.8. To this model, I add variables we might reasonably think are exogenous to BMI status: region of residence, race, and age.[16] Results for this specification are shown in model 2. In models 3-6, I include a different variable (or set of variables) potentially endogenous to obesity; I do this to gauge the importance of each of these co-variates in explaining any observed trends in self-reported health for obese persons. Each of these models includes all of the covariates in model 2 and one of the following potentially endogenous variables (or groups of variables): education (model 3), health conditions (which includes trend-condition, condition-BMI category, and trend-condition-BMI category interactions) (model 4), marital status (model 5), and income (model 6). Finally, model 7 includes all

[15] In order to determine the cutoff for the binary models (excellent, very good, good, fair or poor), I first estimate ordered logit specifications in which I used trends and trend-BMI category interactions as the right-side variables and calculate the marginal effects of these interactions on the probability of all five possible outcomes. In general, I chose the response threshold that seemed to best reflect these marginal effects. I use the binary model to avoid the difficulty of interpreting the ordered models for which one has to examine the relative size and significance of 5 different marginal effects *per variable or interaction* to get a sense of what a given specification yields. Results of ordered models are available from the author on request. For marginal effect of interaction calculations, see Norton et al. (2004) and Ai and Norton (2003).

[16] I use a cubic parameterization of age.

covariates.[17]

In this procedure, the marginal effect of the trend-obesity interaction represents the per-year change in the probability of reporting good or better health for the obese relative to those of normal stature. In general, I expect that any positive marginal effect of this interaction will diminish as covariates related to underlying health–which I capture using the self-reports of the conditions under examination here–enter the model. Model 4 is particularly important in this respect, as it isolates the effect of health conditions as well as condition-BMI category, trend-condition, and trend-condition-BMI category interactions. In essence, while the marginal effects shown in the other columns of Table 3.8 show the per-year change in the probability of good or better health for the obese *with or without* any of the specified conditions relative to those of normal stature with any of them, the marginal effect shown for Model 4 shows the same change in probability for those who do *not* report any of these conditions. If the main hypothesis of this study is true, the latter group–the "healthy obese"–ought to show marginal effects no different from those of the reference group, those of normal stature.[18] Having netted out the effect of health conditions and trends in health conditions with the controls for model 4, I expect that this marginal effect will be smaller in magnitude than those shown in models 1, 2, 3, 5, and 6.

[17]In the discussion of models 4 and 7 that follows, I focus on the trend-, condition-, and trend-condition-obesity interactions. In all cases, models 4 and 7 also include the appropriate interactions with underweight and overweight BMI categories. The reference group is those of normal stature.

[18]I calculate the standard errors of the marginal effects as

$$\hat{\sigma} = \left\{ \frac{\partial \mu}{\partial \beta'} \right\} \hat{\Omega}_{\beta} \left\{ \frac{\partial \mu}{\partial \beta} \right\} + \left\{ n^{-1} * (n-1)^{-1} \sum_{i=1}^{n} (\mu_i - \bar{\mu})^2 \right\}, \tag{3.2}$$

where $\hat{\Omega}_{\beta}$ is a consistent estimate of the covariance matrix of β. The first term on the right hand side of this expression accounts for the variation in the parameter estimates; the second term, frequently overlooked, accounts for the variation in the $X's$. See Basu and Rathouz (2005).

Results

Main Results

Table 3.8 shows the results of this estimation procedure. The top two panels are for women and men in the early waves of NHIS (1982-1996), while the bottom two panels show results for women and men in the later waves (1997-2006).

The results for women in the early waves suggest that there is an improvement in health for obese relative to normal weight women, as was also suggested by the DD estimates. The results from model 1 indicate that, at the end of the 15 years covered by the survey, the probability of reporting good health will have gone up 3.25 percentage points more for obese than normal weight women. This improvement remains even with controls for age, region and race; model 2 implies that the probability of reporting good or better health for obese women will have increased faster than for normal weight women by about 2 percentage points over these years.

For models 3-6, the relevant comparison is between each of the models and the result in model 2 (hereafter, the baseline case), since each model adds only one co-variate (or set of co-variates) to this specification. For women, when controls for education are added (model 3), the measured trend in the baseline model becomes zero. Interestingly, a comparison of model 4 and the baseline model suggests that conditions and their interactions explain very little of this trend: the marginal effect in model 4 is only slightly smaller than in the baseline case, and has similar variance. The results of model 5 suggest that marriage makes obese women feel a little worse about their health, since controlling for it makes the marginal effect of the obesity-trend interaction larger and more significant. As we might expect, income (model 6) has much the same effect as education on the upward trend in health. And model 7, which includes all of the potentially endogenous covariates as well as area level characteristics, suggests that these covariates together explain everything about the upward trend in health that we observe in the baseline case.

The results for men in the earlier waves of the data, shown in the second panel of Table

3.8, suggest that there has been no improvement in health for obese relative to normal weight men over this time period. While the addition of controls for education (model 3), marital status (model 5) and income (model 6) have a modest effect on the estimates, the general message for these models is the same as for the baseline case: there is no evidence that the likelihood of reporting good health changed over these years for men. Interestingly, model 4 suggests that the conditions and condition-trends are negatively correlated with the probability of reporting good health over time. That is, the trends in conditions make it *less* likely that men will report good health over time relative to their normal weight counterparts. We might expect that controlling for conditions alone would have such an effect on the probability of good health because such conditions are negatively correlated with being healthy. However, trends in conditions ought to be positively correlated with trends in health, at least if treatments are related to improving health. The result in Model 4 for men could suggest that any improvements treatments are outweighed by increased severity of conditions. I examine the condition-trend interactions below.

The results for both women and men in the later years of the survey (1997-2006), shown in the bottom two panels of Table 3.8, indicate that there is very little in the way of trends in self-reported health in these years, relative to persons of normal stature. For women, the addition of the controls for health conditions and their interactions (models 4 and 7) have the expected effect: the change in the probability of reporting good health for the "healthy obese" shown in these models is smaller than for the obese as a whole, which suggests that obese persons who reported these conditions did better over time than those without such conditions. For men, these marginal effects are very close to zero in all models. The largest marginal effect (in magnitude) is for model 7, which indicates that health deteriorated more quickly for obese than normal weight men by 5 tenths of one percentage point over these years. This effect is not statistically significant.

Effects of Health Conditions

The hypothesis of this study is that improvements in the treatments for health conditions associated with obesity will have greater beneficial effect on the health of the obese relative to those of normal stature. The results from the top panels of Table 3.8, however, suggest that trends in health for those with these conditions do not explain the upward trend in the probability of reporting good health for obese women, and that they exert a detrimental influence on this probability for obese men, relative to those of normal stature. In order to examine this more closely, I have estimated the marginal effects for each of the conditions studied on its own, interactions of each condition with BMI category indicators (underweight, overweight, obese), interactions of each condition with a linear trend, and 3-way interactions of each condition with a linear trend and BMI indicators. All of these marginal effects are estimated for Model 7 above. The results of this exercise, focused on the obese BMI category, are shown in Tables 3.9 and 3.10 for 1982-1996 and 1997-2006, respectively; the top panels of these Tables show the results for women, while the bottom panels show the results for men.[19]

Since normal weight is the reference category, the marginal effect shown for each condition in the first row of the table is the effect of that condition on persons of normal stature; similarly, the marginal effect of the condition-trend interaction in the second row shows the per year change in the probability of reporting good health for persons of normal stature who report a given condition, relative to this change in probability for those of normal weight who do not report the given condition. The condition-obesity interaction in the third row of the table yields a difference-in-difference marginal effect: it shows the difference in the probability of reporting good health for obese persons with and without the condition relative to normal weight persons with and without the condition. The condition-trend-obesity interaction is a difference-in-difference-in-trends marginal effect: it shows the difference in the per-year change in the probability of reporting good health for obese persons with and

[19] The baseline probability and marginal effects for the linear trend and obesity are shown in Table 3.11 in the Appendix.

without the condition, relative to this same difference in the per-year change in probability for normal weight persons with and without this condition.[20] If the hypothesis of this study is correct, the condition-trend-obesity marginal effects should be positive, as the relative gains in health for obese persons should be greater than for normal weight persons. We would expect this to be true especially if, as seems likely, obese persons are treated more intensively than normal weight persons for the conditions examined here.

The results for women in the earlier waves of the survey suggest that normal weight women with hypertension experience improvements in health over this period, as shown by the positive and significant marginal effect of the condition-trend interaction in the second column. However, none of the marginal effects of condition-trend-obesity interactions are significant, which is consistent with the results shown in model 4 in Table 3.8. There, the controls for condition-trend-obesity interactions do very little to change or explain the estimated marginal effects of the obesity-trend interactions. The results for men in these years (shown in the lower panel of 3.9) are also consistent with what is shown in Table 3.8: the condition-trend-obesity interactions yield negative marginal effects for all conditions except heart attack. That is, as opposed to the "healthy obese" for whom the per-year change in probability of reporting good health is small and positive (Table 3.8, model 7), for obese men with such conditions, the trends are small and negative. In addition, for hypertension and diabetes, the condition-obesity interactions show that obese men who are hypertensive or diabetic are significantly more likely on average to report good health than their normal weight counterparts. Hence, although the trends in the probability of reporting good health are negative, obese hypertensives (diabetics) have a marginal probability of good health 6 (9) percentage points higher than their counterparts without these conditions, relative to normal weight men with and without these conditions. This result is striking because it may indicate that obese men are benefitting from more intensive treatments for these conditions, although improved or intensive treatments do not affect self-reports of health over time.

[20] Alternately, one can think of this three-way interaction as a trend in difference-in-differences.

The results in the top panel of Table 3.10 confirm the results from models 4 and 7 for women in the bottom panel of Table 3.8. Those results suggested that the trends in self-reported health for women who report conditions studied here would be positive, as the marginal effect of the trend-obesity interaction in these models is smaller than in the baseline case. In Table 3.10, all of the condition-trend-obesity marginal effects are positive for women; two are large, and one (for obese women with heart disease) is significant. The marginal effect for heart disease suggests that, although global health improved for obese women with heart disease over these years, the net effect of heart disease for them is still large and negative. The difference in the probability of reporting good health for obese women with and without heart disease relative to the same difference for normal weight women is negative 11.5 percentage points. That is, heart disease affects obese women's likelihood of reporting good health much more strongly than normal weight women's likelihood of reporting good health. While the per-year change in the probability of reporting good health for normal weight women with heart disease is negative 1.5 percentage points, the same change for obese women is positive 1.83 percentage points; the net gain in the per-year probability of reporting good health for obese women is .32 percentage points, which means that by the end of the period examined here, obese women will have closed about 3.2 percentage points of the gap in the probability of reporting good health. Given that the difference-in-difference marginal effect (negative 11.5 percentage points) is an estimate of the *average* difference-in- difference over these years, this improvement represents something less than a third of the difference-in-difference for obese women with heart disease.[21]

The results from the bottom panel of Table 3.10 show that obese male diabetics are at a significant disadvantage on average compared to men of normal stature: the difference-in-difference marginal effect for obese male diabetics (row three of the bottom panel) is negative 10 percentage points. However, the per-year change in the probability of reporting good health for these men is a large and significant positive 2.2 percentage points; the net per-

[21] If the average difference-in-difference gap is -11.5 points, and if the total net increase in the probability of reporting good health is 3.2 percentage points, an estimate of the initial gap would be -11.5-5*(.32) = -13.1 percentage points. Of this gap, 3.2 percentage points is about 24%.

year change relative to normal weight male diabetics (whose per-year change in probability is negative 1 percentage point) is roughly positive 1.2 percentage points. With an average difference-in-difference of negative 10 percentage points, this effect suggests that obese male diabetics closed about 75% of the initial difference-in-difference over these years.[22]

Discussion and Conclusion

Although there is evidence that some of the effects of obesity have been ameliorated by medical innovation in the last 25 years, such relative improvements in health are not uniformly reflected in obese persons' self-assessments of health. The results of this study suggest that, while obese women's health improved in the early waves of the survey relative to their normal weight counterparts, obese men's health did not, once again relative to their normal weight counterparts. The improvements in obese women's health seem particularly tied to education and income and not treatments for health conditions, while for men, the trends belie a significant overall advantage in health for obese male hypertensives and diabetics compared to their normal weight counterparts. The results from the later waves of this study suggest that little is changing in the health of obese women or men over the last 10 years, at least as measured by self-assessments of health. The exceptions to this rule are obese women who report coronary heart disease and obese men who report diabetes: both of these groups report significant improvements health over the later waves of the NHIS.

These results are surprising in several respects. First, for women in the early years of the survey, education is very important to explaining trends in health. While we have known for a long time that education has important–if not quite understood–effects on health, these results suggest that education has very large effects on health trends for obese persons, even given controls for health conditions and condition-trend interactions.

Second, male obese hypertensives and diabetics enjoy a distinct advantage in their health

[22] With an average difference-in-difference of negative 10 percentage points, and an average net gain of 1.2 percentage points for obese male diabetics, a prediction for the difference-in-difference in 1997 would be -10-5*1.2=-16 percentage points; a total net gain of 12 percentage points would represent closing three fourths of this gap.

relative to their normal weight counterparts, as reflected in the results for the later waves of the NHIS. Certainly, it could be that obese men have benefitted on average from more intensive treatments of those conditions. Another explanation for this result has to do with reporting behavior. It could simply be that diabetes and hypertension do less to impair the sense of overall health of obese than normal weight men. In fact, with the exception of heart attack, all of the condition-obese interactions have positive marginal effects during the early years of the survey. Moreover, heart disease, heart attack, and arthritis interactions with obesity show positive marginal effects in the later years, with hypertension essentially being zero. And it is worth noting that, even with the positive marginal effect of the condition-obesity interactions, the net effect of these conditions for obese men is to reduce their probability of reporting good health on average.[23]

Reporting behavior may, in fact, be an important factor in explaining these results. According to one strand in the literature on self-reported health, it is not clear that improvements in technology for treating the co-morbidities of obesity will affect perceptions of health, "true" underlying health notwithstanding. In fact, such an appreciation might seem unlikely: studies have shown that self-assessments of health are often guided by adaptation to conditions and reference bias (See Groot, 2000; Costa-Font and Costa-Font, 2009). The former is frequently understood in the context of a negative health shock. Before a sudden illness or painful injury, people might rate their health as a 7 on a scale of 10. Conditional on surviving such a shock, people will often say that their present health is as good or better than before the shock (i.e. a 7 or better), but revise their estimation of their pre-shock health downward, i.e. to the level of a 5. It isn't hard to imagine treatments being assessed in a similar way, though with the revision operating in the opposite direction: someone with high blood pressure might report health before a new pharmaceutical treatment as a 7, afterward an 8, but afterward also revise their pre-treatment health level to 8 or 9.[24]

[23]Each of the difference-in-difference marginal effects in the third row of the panels showing results for men, where positive, is smaller than the uninteracted marginal effect for each condition, shown in the first row. As shown in Appendix Table 3.11, the marginal effect of obesity is also negative. So the marginal effect of these conditions for obese men is always, on average, negative.

[24]This phenomena is clearly related to the literature on the "hedonic treadmill."(See Wu, 2001) The latter refers to the fact that, on a psychic scale, health is a relative measure: people might assess their own

This would certainly make taking trends in self-reports of health as indicators of underlying health problematic, and could explain non-results in the early years for men and later years for men and women.

If adaptation or reporting behavior more generally explain the lack of results in the later years and the anomalies in the effects of conditions on health for obese men, then what of positive trends in the early years of the survey for obese women? The mechanism by which education has its effect on trends in health for obese persons is unclear. Certainly, for obese women, education is strongly trending over these years, but it's not clear if the effect on health is due to income, information, or some other factor. One might argue that the effect has to do with women's expecations about their health, which trend with education, but the adaptation hypothesis could easily apply to these expecations as it does to underlying health.

The unanswered questions raised by this study open up many avenues for further research. First, this study re-emphasizes the need to understand the mechanisms by which education–and socioeconomic status more generally–affects health. Research on the 1980's and 1990's is important in this respect, because education is trending so strongly during this time–especially for women. Second, the results from the later waves of the survey indicate that treatments for coronary heart disease, which overlap with treatments for hypertension, high cholesterol, and heart attack, may have had a very large and important impact on the health of obese women. Similarly, treatments for diabetes may be behind the very large and significant improvements in health for obese male diabetics in these years. Third, research might examine what self-reports of global health and health conditions measure. Research in this area might take advantage of the fact that NHIS is the sampling frame for the Medical Expenditure Panel Survey and so can be matched to it; thus, one can study the differences in condition reporting patterns over the BMI distribution conditional on objective measures; one could also examine differences in marginal effects of (objective)

health relative to their own health a year earlier, to those with similar health conditions, or to a "worst case scenario" of health given their health conditions. Without some Archimedean point to which self-reports of health might refer it is not clear what they reflect.

conditions on self-reported health over the BMI distribution. All of these would be important first steps to understanding the degree to which it has gotten easier to be obese and how we might know this.

Table 3.1: Effects of Design Change on Summary Measurements, NHIS

	1996	1997	$G_{96-97}(\%)$	$\bar{G}_{82-96}{}^{4}(\%)$	$\bar{G}_{97-06}{}^{4}(\%)$
	Women[2]				
High School Grad	0.3648	0.2939	-0.1944	-0.0133	-0.0134
	(0.4814)	(0.4556)			
Some College[1]	0.2340	0.2898	0.2385	0.0336	0.0033
	(0.4234)	(0.4537)			
Income Percentile	0.5081	0.5641	0.1102	0.0032	0.0035
	(0.3742)	(0.3075)			
Very Good Health	0.2942	0.3255	0.1063	0.0124	-0.0025
	(0.4557)	(0.4686)			
	Men[3]				
High School Grad	0.3320	0.2748	-0.1724	0.0023	0.0031
	(0.4710)	(0.4464)			
Some College[1]	0.2227	0.2797	0.2563	0.0232	-0.0120
	(0.4161)	(0.4489)			
Income Percentile	0.5641	0.6163	0.0926	0.0009	0.0004
	(0.3713)	(0.2960)			
Very Good Health	0.3007	0.3282	0.0916	0.0156	0.0078
	(0.4586)	(0.4696)			

[1] Indicates having some college education but no college degree.

[2] Sample sizes for women in 1996 and 1997 are 4,909 and 6,496, respectively.

[3] Sample sizes for men in 1996 and 1997 are 7,410 and 7,660, respectively.

[4] Calculated as the geometric mean rate of growth over survey means by year.

Table 3.2: Self-Reports of Health Conditions in NHIS

Condition	Survey Question
1982-1996	
Arthritis	During the past 12 months, did you have arthritis?
Diabetes	During the past 12 months, did you have diabetes?
Hypertension	Have you ever had hypertension?
Coronary Heart Disease	Have you ever had coronary heart disease?
Heart Attack	Have you ever had a heart attack?
All	What condition causes [...] limitation of activity?
	What condition caused you to [miss work, miss school, stay in bed, cut down your activity]?
	For what condition did you see/talk to a doctor/medical specialist?
	For what condition did you enter the hospital?
1997-2006	
Arthritis	Are you limited in any way in any activities by physical, mental, or emotional problems?
	What condition causes this limitation?
Diabetes	Have you ever been told by a health professional that you had diabetes?
Hypertension	Have you ever been told by a health professional that you had hypertension?
Coronary Heart Disease	Have you ever been told by a health professional that you had coronary heart disease?
Heart Attack	Have you ever been told by a health professional that you had a heart attack?

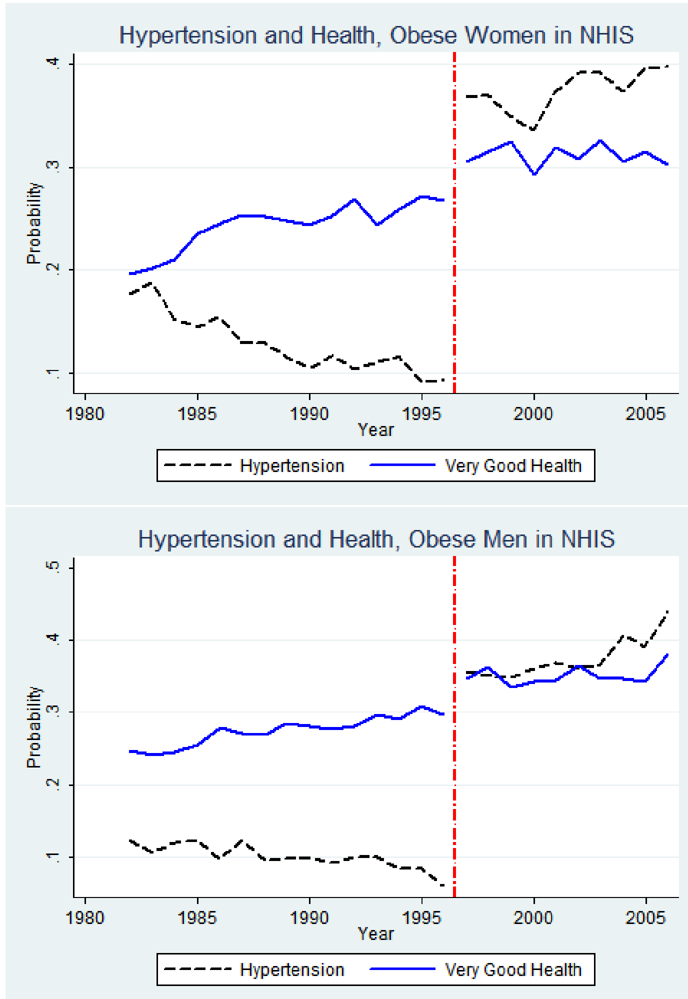


Figure 3.1: Hypertension and Health, Obese Women and Men in NHIS

Table 3.3: Summary Statistics: NHIS

White	0.792	0.742	0.795	0.756
	(0.406)	(0.438)	(0.403)	(0.429)
Black	0.111	0.116	0.101	0.100
	(0.314)	(0.321)	(0.302)	(0.300)
Hispanic	0.068	0.101	0.070	0.103
	(0.252)	(0.301)	(0.255)	(0.304)
Other Race	0.029	0.041	0.034	0.041
	(0.166)	(0.198)	(0.180)	(0.197)
Age	45.946	46.362	46.044	46.258
	(7.572)	(7.207)	(7.563)	(7.174)
North Central Region	0.255	0.241	0.247	0.249
	(0.436)	(0.428)	(0.431)	(0.433)
West Region	0.203	0.198	0.227	0.201
	(0.402)	(0.398)	(0.419)	(0.401)
Northeast Region	0.205	0.204	0.204	0.191
	(0.404)	(0.403)	(0.403)	(0.393)
More Than HS Education	0.393	0.577	0.481	0.574
	(0.488)	(0.494)	(0.500)	(0.495)
Married	0.702	0.679	0.699	0.718
	(0.457)	(0.467)	(0.459)	(0.450)
Excellent Health	0.317	0.307	0.356	0.324
	(0.465)	(0.461)	(0.479)	(0.468)
Very Good Health	0.281	0.334	0.278	0.341
	(0.449)	(0.471)	(0.448)	(0.474)
Good Health	0.267	0.247	0.234	0.238
	(0.442)	(0.431)	(0.424)	(0.426)
Fair Health	0.098	0.084	0.085	0.072
	(0.297)	(0.277)	(0.279)	(0.259)
Poor Health	0.038	0.028	0.047	0.025
	(0.190)	(0.166)	(0.212)	(0.156)
Underweight	0.024	0.012	0.007	0.001
	(0.153)	(0.110)	(0.082)	(0.039)
Normal Weight	0.483	0.397	0.370	0.268
	(0.500)	(0.489)	(0.483)	(0.443)
Overweight	0.284	0.309	0.443	0.464
	(0.451)	(0.462)	(0.497)	(0.499)
Obese	0.209	0.282	0.180	0.266
	(0.406)	(0.450)	(0.384)	(0.442)
Hypertension	0.057	0.216	0.054	0.239
	(0.233)	(0.412)	(0.226)	(0.426)
Diabetes	0.022	0.053	0.021	0.058
	(0.145)	(0.224)	(0.144)	(0.234)
Heart Attack	0.003	0.012	0.011	0.028
	(0.058)	(0.108)	(0.106)	(0.165)
Coronary Heart Disease	0.009	0.069	0.022	0.051
	(0.095)	(0.253)	(0.147)	(0.220)
Arthritis	0.062	0.028	0.041	0.015
	(0.240)	(0.164)	(0.198)	(0.122)

Continued on Next Page

Table 3.3: Summary Statistics: NHIS, continued

	Women		Men	
Variable	1982-1996	1997-2006	1982-1996	1997-2006
Number of Conditions Reported	1.029	0.217	0.874	0.189
	(1.460)	(0.759)	(1.344)	(0.683)
Area Health	3.724	3.804	3.734	3.813
	(0.215)	(0.287)	(0.209)	(0.284)
Area BMI	25.642	27.130	25.653	27.135
	(0.811)	(1.176)	(0.805)	(1.148)
N	183,947	67,931	107,367	58,389

Table 3.4: Probability of Health Conditions by BMI Status, NHIS 1982-1996[1]

	Underweight	Normal Weight	Overweight	Obese
	Women[2]			
Hypertension	0.029	0.027	0.061	0.126
	(0.168)	(0.161)	(0.239)	(0.338)
Diabetes	0.011	0.007	0.019	0.059
	(0.103)	(0.085)	(0.136)	(0.240)
Heart Attack	0.005	0.002	0.004	0.006
	(0.068)	(0.045)	(0.061)	(0.077)
Coronary Heart Disease	0.009	0.005	0.010	0.016
	(0.095)	(0.071)	(0.101)	(0.129)
Arthritis	0.051	0.040	0.063	0.111
	(0.220)	(0.195)	(0.243)	(0.319)
	Men[3]			
Hypertension	0.053	0.033	0.053	0.098
	(0.228)	(0.179)	(0.224)	(0.297)
Diabetes	0.026	0.014	0.017	0.044
	(0.162)	(0.119)	(0.131)	(0.204)
Heart Attack	0.008	0.008	0.012	0.017
	(0.091)	(0.091)	(0.107)	(0.128)
Coronary Heart Disease	0.021	0.016	0.023	0.034
	(0.147)	(0.125)	(0.149)	(0.180)
Arthritis	0.065	0.033	0.039	0.061
	(0.251)	(0.180)	(0.193)	(0.238)

[1] Estimates are computed using survey design variables. Standard deviation in parenthesis.

[2] Sample size for women in 183,947.

[3] Sample size for men is 107,367.

Table 3.5: Probability of Health Conditions by BMI Status, NHIS 1997-2006[1]

	Underweight	Normal Weight	Overweight	Obese
	Women[2]			
Hypertension	0.086	0.111	0.211	0.376
	(0.274)	(0.308)	(0.410)	(0.496)
Diabetes	0.029	0.018	0.043	0.116
	(0.164)	(0.128)	(0.203)	(0.328)
Heart Attack	0.006	0.006	0.011	0.020
	(0.074)	(0.078)	(0.103)	(0.145)
Coronary Heart Disease	0.077	0.063	0.065	0.081
	(0.260)	(0.237)	(0.247)	(0.280)
Arthritis	0.016	0.013	0.025	0.051
	(0.123)	(0.113)	(0.157)	(0.226)
	Men[3]			
Hypertension	0.193	0.141	0.216	0.378
	(0.408)	(0.355)	(0.411)	(0.476)
Diabetes	0.081	0.031	0.044	0.110
	(0.283)	(0.177)	(0.205)	(0.308)
Heart Attack	0.061	0.020	0.025	0.041
	(0.247)	(0.142)	(0.156)	(0.195)
Coronary Heart Disease	0.087	0.045	0.048	0.062
	(0.292)	(0.211)	(0.213)	(0.236)
Arthritis	0.038	0.013	0.011	0.025
	(0.197)	(0.115)	(0.104)	(0.152)

[1] Estimates are computed using survey design variables. Standard deviation in parenthesis.

[2] Sample size for women in 67,931.

[3] Sample size for men is 58,389.

Table 3.6: Difference In Differences for Early Waves of NHIS[1]

	1982			1996			
	Normal Wgt_{82}[1]	$Obese_{82}$	$Diff_{82}$[2]	Normal Wgt_{96}	$Obese_{96}$	$Diff_{96}$	$D\text{-}D_{96-82}$[3]
	Women[4]						
Very Good or Excellent Health	0.6558	0.3557	0.3000	0.7120	0.4438	0.2682	-0.0318
	(0.0056)	(0.0113)	(0.0125)	(0.0122)	(0.0174)	(0.0179)	(0.0218)
Fair or Poor Health	0.1061	0.2945	-0.1884	0.0824	0.2373	-0.1549	0.0336*
	(0.0040)	(0.0107)	(0.0114)	(0.0078)	(0.0149)	(0.0145)	(0.0184)
	Men[5]						
Very Good or Excellent Health	0.6180	0.4719	0.1462	0.7028	0.5108	0.1921	0.0459
	(0.0110)	(0.0153)	(0.0204)	(0.0137)	(0.0220)	(0.0264)	(0.0334)
Fair or Poor Health	0.1562	0.2327	-0.0765	0.1084	0.1786	-0.0702	0.0064
	(0.0071)	(0.0137)	(0.0152)	(0.0138)	(0.0198)	(0.0222)	(0.0269)

[1] Estimates are computed using survey design variables. Standard errors in parenthesis.

[2] All differences are Normal Wgt-Obese.

[3] ***p<.01, **p<.05, *p<.10

[4] Samples sizes for women in years 1982 and 1996 are 11,804 and 7,410, respectively.

[5] Samples sizes for men in years 1982 and 1996 are 6,406 and 4,909, respectively.

Table 3.7: Difference In Differences for Later Waves of NHIS[1]

	1997			2006			
	Normal Wgt_{97}[1]	$Obese_{97}$	$Diff_{97}$[2]	Normal Wgt_{06}	$Obese_{06}$	$Diff_{06}$	$D\text{-}D_{06-97}$[3]
	Women[4]						
Very Good or Excellent Health	0.7502	0.4749	0.2752	0.7143	0.4537	0.2607	-0.0146
	(0.0084)	(0.0133)	(0.0157)	(0.0126)	(0.0148)	(0.0195)	(0.0250)
Fair or Poor Health	0.0597	0.1869	-0.1272	0.0737	0.1928	-0.1190	0.0082
	(0.0047)	(0.0099)	(0.0105)	(0.0068)	(0.0117)	(0.0133)	(0.0170)
	Men[4]						
Very Good or Excellent Health	0.6939	0.5837	0.1103	0.6843	0.5656	0.1187	0.0084
	(0.0116)	(0.0144)	(0.0171)	(0.0169)	(0.0145)	(0.0225)	(0.0283)
Fair or Poor Health	0.0973	0.1336	-0.0363	0.1002	0.1376	-0.0374	-0.0011
	(0.0067)	(0.0110)	(0.0128)	(0.0095)	(0.0093)	(0.0133)	(0.0185)

[1] Estimates are computed using survey design variables. Standard errors in parenthesis.

[2] All differences are Normal Wgt-Obese.

[3] p<.01, **p<.05, *p<.10

[4] Samples sizes for women in years 1997 and 2006 are 7,778 and 5,384, respectively.

[5] Samples sizes for men in years 1997 and 2006 are 6,602 and 4,738, respectively.

Table 3.8: Marginal Effects on Health, Women and Men in NHIS[1]

	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7
	1982-1996[2]						
	Women						
Obese×Trend	0.00217*** (0.00076)	0.00135* (0.00070)	0.00022 (0.00065)	0.00125* (0.00069)	0.00153** (0.00070)	0.00087 (0.00064)	-0.00007 (0.00056)
	Men						
Obese×Trend	0.00034 (0.00094)	0.00045 (0.00087)	-0.00049 (0.00080)	0.00165** (0.00083)	0.00051 (0.00088)	-0.00032 (0.00084)	0.00035 (0.00072)
	1997-2006[3]						
	Women						
Obese×Trend	-0.00086 (0.00136)	-0.00008 (0.00124)	-0.00036 (0.00116)	-0.00229 (0.00160)	-0.00011 (0.00122)	-0.00017 (0.00105)	-0.00180 (0.00131)
	1997-2006						
	Men						
Obese×Trend	-0.00016 (0.00125)	0.00011 (0.00114)	-0.00044 (0.00105)	-0.00052 (0.00144)	0.00029 (0.00113)	-0.00033 (0.00099)	-0.00053 (0.00127)
Trend, G^4×Trend	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Age, Region, Race	No	Yes	Yes	Yes	Yes	Yes	Yes
Education	No	No	Yes	No	No	No	Yes
C^5, C×Trend,C×G×Trend	No	No	No	Yes	No	No	Yes
Marital Status	No	No	No	No	Yes	No	Yes
Income	No	No	No	No	No	Yes	Yes
Area Variables	No	No	No	No	No	No	Yes

[1] Marginal effects calculated from results of binary logit models. Dependent variable is an indicator of good, very good, or excellent health. Standard errors in parenthesis. *** p<.01, **p<.05, *p<.10.

[2] Sample sizes for women and men are 183,947 and 107,367, respectively.

[3] Sample sizes for women and men are 67,931 and 58,389, respectively.

[4] G denotes set of BMI-category dummies: underweight, overweight, obese. Normal weight is the reference category.

[5] C denotes indicators for hypertension, coronary heart disease, heart attack, diabetes, and arthritis. This model also includes the number of conditions reported other than those listed here.

Table 3.9: Marginal Effect of Health Conditions on Global Health, NHIS 1982-1996[1]

	Heart Disease	Hypertension	Heart Attack	Diabetes	Arthritis
			Women[2]		
Cond	-0.1403***	-0.1085***	0.0027	-0.1300***	-0.0640***
	(0.0269)	(0.0097)	(0.0440)	(0.0173)	(0.0092)
Cond×Trend	-0.0005	0.0044**	0.0006	-0.0011	0.0005
	(0.0073)	(0.0020)	(0.0053)	(0.0041)	(0.0016)
Cond×Obese	-0.0079	0.0148	0.0236	-0.0026	-0.0244
	(0.0843)	(0.0207)	(0.0686)	(0.0393)	(0.0190)
Cond×Trend×Obese	0.0069	-0.0001	-0.0080	0.0019	-0.0002
	(0.0100)	(0.0024)	(0.0099)	(0.0047)	(0.0022)
			Men[3]		
Cond	-0.1687***	-0.1040***	0.0112	-0.1341***	-0.0638***
	(0.0235)	(0.0118)	(0.0318)	(0.0220)	(0.0129)
Cond×Trend	0.0073	0.0075***	-0.0043	0.0043	0.0004
	(0.0063)	(0.0022)	(0.0051)	(0.0047)	(0.0023)
Cond×Obese	0.0799	0.0578**	-0.0454	0.0928*	0.0126
	(0.0832)	(0.0281)	(0.0659)	(0.0559)	(0.0304)
Cond×Trend×Obese	-0.0106	-0.0052	0.0085	-0.0091	-0.0001
	(0.0094)	(0.0032)	(0.0068)	(0.0061)	(0.0037)

[1] Marginal effects calculated from results of a binary logit model. Dependent variable is an indicator of good, very good, or excellent health. Results are for model 7 in Table 3.8, which includes G, a set of BMI category dummies (underweight, overweight, obese) for which normal weight is the reference category; a linear trend and the interactions between G and the trend; race, age, and region; education, marital status, income, area-level health, and area-level BMI; health conditions listed in the table, condition interactions with the linear time trend, and condition-trend-G interactions. Standard errors in parenthesis. *** p<.01, **p<.05, *p<.10.

[2] Sample size for women is 183,947.

[3] Sample size for men is 107,367.

Table 3.10: Marginal Effect of Health Conditions on Global Health, NHIS 1997-2006[1]

	Heart Disease	Hypertension	Heart Attack	Diabetes	Arthritis
	Women[2]				
Cond	-0.0864***	-0.1200***	-0.0645*	-0.1184***	-0.0835***
	(0.0168)	(0.0107)	(0.0378)	(0.0256)	(0.0262)
Cond×Trend	-0.0151***	0.0003	-0.0048	-0.0161	-0.0143
	(0.0055)	(0.0033)	(0.0103)	(0.0101)	(0.0100)
Cond×Obese	-0.1147***	-0.0385	-0.0188	-0.0991	0.0450
	(0.0422)	(0.0255)	(0.0764)	(0.0644)	(0.0582)
Cond×Trend×Obese	0.0183**	0.0012	0.0080	0.0120	0.0092
	(0.0072)	(0.0039)	(0.0122)	(0.0112)	(0.0114)
	Men[2]				
Cond	-0.1202***	-0.0947***	-0.0469**	-0.1237***	-0.1048***
	(0.0151)	(0.0110)	(0.0224)	(0.0179)	(0.0286)
Cond×Trend	0.0034	-0.0021	-0.0034	-0.0102	0.0098
	(0.0064)	(0.0035)	(0.0061)	(0.0083)	(0.0095)
Cond×Obese	0.0642	-0.0030	0.0033	-0.1034	0.0742
	(0.0539)	(0.0251)	(0.0439)	(0.0641)	(0.0816)
Cond×Trend×Obese	-0.0049	-0.0009	0.0020	0.0222**	-0.0070
	(0.0086)	(0.0042)	(0.0073)	(0.0102)	(0.0119)

[1] Marginal effects calculated from results of a binary logit model. Dependent variable is an indicator of good, very good, or excellent health. Results are for model 7 in Table 3.8, which includes G, a set of BMI category dummies (underweight, overweight, obese) for which normal weight is the reference category; a linear trend and the interactions between G and the trend; race, age, and region; education, marital status, income, area-level health, area-level BMI, smoking and exercise; health conditions listed in the table, condition interactions with the linear time trend, and condition-trend-G interactions. Standard errors in parenthesis. *** p<.01, **p<.05, *p<.10.

[2] Sample size for women is 67,034.

[3] Sample size for men is 57,324.

Appendix A: Formulas for 3-way Interaction Marginal Effects

Formulas for 2-way interactions can be found in Norton et al. (2004) and Ai and Norton (2003). For 3-way interactions, let $x_1 = I * (BMI \geq 30)$, $x_2 = I * (HealthCond = 1)$, and x_3 =linear trend, where $I = 1$ if stated condition is true, and $HealthCond = 1$ indicates that a respondent says that a given health condition is present.

$$P(H_i = 1) = F(X\beta + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_{12} x_1 x_2 + \beta_{13} x_1 x_3 + \beta_{23} x_2 x_3 + \beta_{123} x_1 x_2 x_3), \quad (3.3)$$

where $F()$ is the logistic CDF and X is comprised of other covariates in the model. Let

$$\begin{aligned} \eta_1 &= X\beta + \beta_1 + \beta_2 + \beta_3 x_3 + \beta_{12} + \beta_{13} x_3 + \beta_{23} x_3 + \beta_{123} x_3 \\ \eta_2 &= X\beta + \beta_2 + \beta_3 x_3 + \beta_{23} x_3 \\ \eta_3 &= X\beta + \beta_1 + \beta_3 x_3 + \beta_{13} x_3 \\ \eta_4 &= X\beta + \beta_3 x_3. \end{aligned} \quad (3.4)$$

Then

$$\frac{\partial\{\frac{\Delta^2 P(H_i=1)}{\Delta x_1 \Delta x_2}\}}{\partial x_3} = \{(\beta_3+\beta_{13}+\beta_{23}+\beta_{123}) * F'(\eta_1) - (\beta_3+\beta_{23}) * F'(\eta_2)\} - \{(\beta_3+\beta_{13}) * F'(\eta_3) - \beta_3 * F'(\eta_4)\}, \quad (3.5)$$

where $F'()$ is the logistic PDF. I calculate marginal effects for each person. The values shown throughout are the mean of those marginal effects. As mentioned above, I calculate the standard errors of the marginal effects as

$$\hat{\sigma} = \left\{\frac{\partial \mu}{\partial \beta'}\right\} \hat{\Omega}_\beta \left\{\frac{\partial \mu}{\partial \beta}\right\} + \left\{n^{-1} * (n-1)^{-1} \sum_{i=1}^{n} (\mu_i - \bar{\mu})^2\right\}, \quad (3.6)$$

where $\mu = \frac{\partial\{\frac{\Delta^2 P(H_i=1)}{\Delta x_1 \Delta x_2}\}}{\partial x_3}$.

Appendix B: Probabilities and Marginal Effects, Fully Conditional Model

Table 3.11: Probability and Marginal Effects, Model 7, NHIS[1]

	1982-1996		1997-2006	
	Female	Male	Female	Male
Constant[2]	0.8650***	0.8689***	0.8748***	0.8859***
	(0.0037)	(0.0056)	(0.0099)	(0.0115)
Obese	-0.0470***	-0.0217***	-0.0280***	-0.0090***
	(0.0018)	(0.0011)	(0.0025)	(0.0009)
Trend	-0.0006	0.0006	-0.0015*	-0.0002
	(0.0004)	(0.0005)	(0.0008)	(0.0009)
N	183,947	107,367	67,931	58,389

[1] Marginal effects calculated from results of binary logit models. Dependent variable is an indicator of good, very good, or excellent health. Standard errors in parenthesis. *** p<.01, **p<.05, *p<.10.

[2] Reference group is normal weight white persons who live in the South, are either high school graduates or high school dropouts, and do not report hypertension, diabetes, coronary heart disease, heart attack, or arthritis as health conditions. In the later waves of the survey, the reference group are, in addition, non-smokers.

Bibliography

Ai, C. and Norton, E. C. (2003). Interaction effects in logit and probit models. *Economics Letters*, 80:123–129.

Andreyeva, T., Sturm, R., and Ringel, J. S. (2004). Moderate and severe obesity have large differences in health costs. *Obesity Research*, 12:1936–1943.

Arterburn, D., Maciejewski, M., and Tsevat, J. (2005). Impact of morbid obesity on medical expenditures in adults. *International Journal of Obesity*, 29:324–339.

Averett, S. and Korenmann, S. (1996). The economic reality of the beauty myth. *The Journal of Human Resources*, 31:304–330.

Baker, M., Stabile, M., and Deri, C. (2004). What do self-reported, objective, measures of health measure? *Journal of Human Resources*, 39(4):1067–1093.

Basu, A. and Rathouz, P. J. (2005). Estimating marginal and incremental effects on health outcomes using flexible link and variance function models. *Biostatistics*, 6(1):93–109.

Baum, C. and Ruhm, C. J. (2007). Age, socioeconomic status, and obesity growth. Technical Report 13289, NBER Working Paper.

Baum, C. L. and Ford, W. F. (2004). The wage effects of obesity: A longitudinal study. *Health Economics*, 13:885–899.

Becker, H. F., Jerrentrup, A., Ploch, T., Grote, L., Penzel, T., Sullivan, C. E., and Peter, J. H. (2003). Effect of nasal continuous positive airway pressure treatment on blood pressure in patients with obstructive sleep apnea. *Circulation*, 107(1):68–73.

Bhattacharya, J. and Bundorf, K. (2005). The incidence of the healthcare costs of obesity. *Journal of Health Economics*, 28(3):649–658.

Biddle, J. E. and Hamermesh, D. S. (1994). Beauty, discrimination and the labor market. *American Economic Review*, 84:1174–1194.

Bollinger, C. R. and Chandra, A. (2005). Iatrogenic specification error: A cautionary tale of cleaning data. *Journal of Labor Economics*, 23(2):235–257.

Bound, J., Brown, C., Duncan, G. J., and Rodgers, W. L. (1994). Evidence on the validity of cross-sectional and longitudinal labor market data. *Journal of Labor Economics*, 12:345–368.

Bound, J., Brown, C., and Mathiowetz, N. (2001). Measurement error in survey data. In Heckman, J. and Learner, E., editors, *The Handbook of Econometrics*, volume 5, pages 3707–3843. Elsevier.

Burkhauser, R. V. and Cawley, J. (2008a). Beyond BMI: The value of more accurate measures of fatness and obesity in social science research. *Journal of Health Economics*, 27:519–529.

Burkhauser, R. V. and Cawley, J. (2008b). Beyond BMI: The value of more accurate measures of fatness and obesity in social science research. *Journal of Health Economics*, 27(2):519–529.

Burt, V. L., Cutler, J. A., Higgins, M., Horan, M. J., Labarthe, D., Whelton, P., Brown, C., and Roccella, E. J. (1995). Trends in the prevalence, awareness, treatment, and control of hypertension in the adult us population: Data from the health examination surveys 1960 to 1991. *Hypertension*, 26(1):60–69.

Cawley, J. (2004). The impact of obesity on wages. *Journal of Human Resources*, 39:451–474.

Chou, S.-Y., Grossman, M., and Saffer, H. (2004). An economic analysis of adult obesity: Results from the behavioral risk factor surveillance system. *Journal of Health Economics*, 23:565–587.

Costa-Font, M. and Costa-Font, J. (2009). Heterogenous 'adaptation' and 'income effects' across self-reported health distribution. *Journal of Socio-Economics*, 38(4):574–580.

Craven, P. and Wahba, G. (1979). Smoothing noisy data with spline functions. *Numerische Mathematik*, 13:377–403.

Cutler, D., Glaeser, E., and Shapiro, J. M. (2003). Why have Americans become more obese? *Journal of Economic Perspectives*, 17:93–118.

Cutler, J. A., Sorlie, P. D., Wolz, M., Thom, T., Fields, L. E., and Roccella, E. J. (2008). Trends in the prevalence, awareness, treatment, and control of hypertension in us adults between 1988-1994 and 1999-2004. *Hypertension*, 52(5):818–827.

Daniels, M. J. and Hogan, J. W. (2008). *Missing Data in Longitudinal Studies: Strategies for Bayesian Modeling and Sensitivity Analysis*. Chapman & Hall/CRC.

Deurenberg, P., Westrate, J. A., and Seidell, J. C. (1991). Body mass index as a measure of body fatness: age- and sex- specific prediction formulas. *British Journal of Nutrition*, 65(2):105–144.

Diabetes Prevention Program Research Group (2002). Reduction in the incidence of type 2 diabetes with lifestyle intervention or metformin. *New England Journal of Medicine*, 346(6):393–403.

DiGioachino, R. F., Sargent, R., and Topping, M. (2001). Body dissatisfaction among white and african american male and female college students. *Eating Behaviors*, 2:39–50.

Dong, C., Sanchez, L. E., and Price, R. A. (2004). Relationship of obesity to depression: A family based study. *International Journal of Obesity*, 28:790–795.

Duncan, G. J. and Hill, D. H. (1985). An investigation of the extent and consequences of measurement error in labor-economic survey data. *Journal of Labor Economics*, 3:508–532.

El-Faddagh, M., Laucht, M., Maras, A., Vohringer, L., and Schmidt, M. (2004). Association of dopamine d4 receptor gene with attention-deficity/hyperactivity disorder in a high risk community sample: a longitudinal study from birth to 11 years of age. *Journal of Neural Transmission*, 111:883–889.

Endo, A. (1992). The discovery and development of HMG-CoA reductase inhibitors. *Journal of Lipid Research*, 33:1569–1582.

Fan, J. and Gijbels, I. (1996). *Local Polynomial Modelling and Its Applications*. CRC Press.

Finkelstein, E. A., Brown, D. S., Trogdon, J. G., Segel, J. E., and Ben-Joseph, R. H. (2007). Age-specific impact of obesity on prevalence and costs of diabetes and dyslipidemia. *Value in Health*, 10:S45–S51.

Flegal, K. M., Carroll, M. D., Ogden, C. L., and Curtin, L. R. (2010). Prevalence and trends in obesity among us adults 1999-2008. *JAMA*, 303(3):235–241.

Flegal, K. M., Graubard, B. I., Williamson, D. F., and Gail, M. H. (2005). Excess deaths associated with underweight, overweight, and obesity. *JAMA*, 293:1861–1867.

Flegal, K. M., Graubard, B. I., Williamson, D. F., and Gail, M. H. (2007). Excess deaths associated with underweight, overweight, and obesity. *JAMA*, 298:2028–2037.

Folli, F., Pontiroli, A. E., and Schwesinger, W. H. (2007). Metabolic aspects of bariatric surgery. *Medical Clinics of North America*, 91:393–414.

Ford, E. S., Moriarty, D. G., Zack, M. M., Mokdad, A. H., and Chapman, D. P. (2001). Self-reported body mass index and health-related quality of life: Findings from the behavioral risk factor surveillance system. *Obesity Research*, 9(1):21–31.

French, M. T. (2002). Physical appearance and earnings: Further evidence. *Applied Economics*, 34:569–572.

Gortmaker, S. L., Must, A., and Perrin, J. M. (1993). Social and economics consequences of overweight in adolescence and young adulthood. *New England Journal of Medicine*, 329(14):1008–1012.

Gregg, E. W., Cheng, Y. J., Cadwell, B. L., Imperatore, G., Williams, D. E., Flegal, K. M., Narayan, K. M. V., and Williamson, D. F. (2005). Secular trends in cardiovascular disease risk factors according to body mass index in us adults. *JAMA*, 293:1868–1874.

Gregg, E. W. and Guralnik, J. M. (2007). Is disability obesity's price of longevity? *JAMA*, 298:2066–2067.

Gregory, C. A. (2010). Wages, age, and BMI. Unpublished MS, University of North Carolina at Greensboro.

Gregory, C. A. and Ruhm, C. J. (2009). Where does the wage penalty bite? Working Paper 14984, National Bureau of Economic Research.

Groot, W. (2000). Adaptation and scale of reference bias in self-assesments of quality of life. *Journal of Health Economics*, 19(3):403–420.

Grossman, M. (2000). The human capital model. In Culyer, J. and Newhouse, J. P., editors, *The Handbook of Health Economics*, volume 1A, pages 347–408. Elsevier.

Hamermesh, D. S. and Biddle, J. E. (1998). Beauty, productivity and discrimination: Lawyers' looks and lucre. *Journal of Labor Economics*, 16:172–201.

Han, E., Norton, E. C., and Powell, L. M. (2009a). Direct and indirect effects of teenage body weight on adult wages. NBER Working Papers 15027, National Bureau of Economic Research, Inc.

Han, E., Norton, E. C., and Stearns, S. C. (2009b). Weight and wages: Fat versus lean paychecks. *Health Economics*, 18(5):535–548.

Härdle, W. (1990). *Applied Nonparametric Regression*. Cambridge University Press.

Harper, B. (2000). Beauty, stature and the labour market: A british cohort study. *Oxford Bulletin of Economics and Statistics*, 62:771–800.

Heckman, J. J. (1979). Sample selection bias as a specification error. *Econometrica*, 47(1):153–161.

Hill, M. S. (1992). *The Panel Study of Income Dynamics: A User's Guide*. Sage Publications.

Holtkamp, K., Konrad, K., Muller, B., Heussen, N., Herpertz, S., Herpertz-Dahlmann, B., and Hebebrand, J. (2004). Overweight and obesity in children with attention-deficit/hyperactivity disorder. *International Journal of Obesity*, 28:685–689.

Idler, E. L. and Kasl, S. V. (1995). Self-ratings of health: Do they also predict change in functional ability? *Journal of Gerontology*, 50B(6):S344–S353.

Johannson, E., Bockerman, P., Kiiskinen, U., and Heliövaara, M. (2009). Obesity and labour market success in finland: the difference between having a high BMI and being fat. *Economics and Human Biology*, 7(1):36–45.

Johnston, D. W., Propper, C., and Sheilds, M. A. (2009). Comparing subjective and objective measures of health: Evidence from hypertension for the income/health gradient. *Journal of Health Economics*, 28(3):540–552.

Kaplan, G. A., Strawbridge, W. J., Camacho, T., and Cohen, R. (1993). Factors associated with change in physical functioning in the elderly. *Journal of Aging and Health*, 5(1):140–153.

Kawachi, I., Kennedy, B. P., and Glass, R. (1989). Social capital and self-rated health: A contextual analysis. *American Journal of Public Health*, 89(8):1187–1193.

Kline, B. and Tobias, J. L. (2008). The wages of BMI: Bayesian analysis of a skewed treatment-response model with non-parametric endogeneity. *Journal of Applied Econometrics*, 23(6):767–793.

Krause, N. M. and Jay, G. M. (1994). What do global self-rated health items measure? *Medical Care*, 32(9):930–942.

Lakdawalla, D. and Philipson, T. (2007). Labor supply and body weight. *Journal of Human Resources*, 42:85–116.

Lakdawalla, D. and Phillipson, T. J. (2002). Technological change and the growth of obesity. *NBER Working Paper*, (8946).

Lakdawalla, D. and Phillipson, T. J. (2007). Labor supply and body weight. *Journal of Human Resources*, 42:85–116.

Lee, L.-F. and Sepanski, J. H. (1995). Estimation of linear and non-linear errors-in-variables models using validation data. *Journal of the American Statistical Association*, 429:130–140.

Levitan, R., Masells, M., Lam, R., Basile, V., Jain, U., Kaplan, A. S., Thramalingam, S., Kennedy, S., and Kennedy, J. (2004). Childhood inattention and dysphoria and adult obesity associated with the dopamine d4 receptor gene in overeating women with seasonal affective disorder. *Neuropsychopharmacology*, 29:179–196.

Li, Q. and Racine, J. S. (2007). *Nonparametric Econometrics*. Princeton UP.

Loh, E. S. (1993). The economic effects of physical appearance. *Social Science Quarterly*, 74(2):420–438.

MacCurdy, T., Mroz, T., and Gritz, R. (1998). An evaluation of the National Longitudinal Survey of Youth. *Journal of Human Resources*, 33(2):345–436.

Maes, H. H. M., Neale, M. C., and Eaves, L. J. (1997). Genetic and environmental factors in relative body weight and human adiposity. *Behavior Genetics*, 27:325–351.

Marmot, M., Feeney, A., Shipley, M., North, F., and Syme, S. (1995). Sickness absence as a measure of health status and functioning: from the uk whitehall ii study. *Journal of Epidemiology and Community Health*, 49:124–130.

Maynard, L., Serdula, M., Galuska, D., Gillespie, C., and Mokdad, A. (2006). Secular trends in desired weight of adults. *International Journal of Obesity*, 30:1375–1381.

McGee, D. L., Liao, Y., Guichan, and Cooper, R. S. (1999). Self-reported health status and mortality in a multiethnic us cohort. *American Journal of Epidemiology*, 149(1):41–46.

Mello, M. M., Studdert, D. M., and Brennan, T. A. (2006). Obesity–the new frontier of public health law. *New England Journal of Medicine*, 354:2601–2610.

Miilunpalo, S., Vuori, I., Oja, P., Pasanen, M., and Urponen, H. (1997). Self-rated health status as a health measure: The predictive value of self-reported health status on the use of physician services and on mortality in the working age population. *Journal of Clinical Epidemiology*, 50(5):517–528.

Mobius, M. M. and Rosenblat, T. S. (2006). Why beauty matters. *American Economic Review*, 96:222–235.

Mocan, N. and Tekin, E. (2006). Ugly criminals. *NBER Working Paper*, (12019).

Müller, M. (2001). Estimation and testing in generalized linear models–a comparative study. *Statistics and Computing*, 11:299–309.

Must, A., Spadano, J., Coakley, E. H., Field, A. E., Colditz, G., and Dietz, W. H. (1999). The disease burden associated with overweight and obesity. *JAMA*, 282:1523–1529.

National Heart, Lung and Blood Institute (1998). *Clinical Guidelines on the Identification, Evaluation, and Treatment of Overweight and Obesity in Adults: The Evidence Report*. U.S. Government Priting Office.

Norton, E. C. and Han, E. (2008). Genetic information, obesity, and labor market outcomes. *Health Economics*, 17(9):1089–1104.

Norton, E. C., Wang, H., and Ai, C. (2004). Computing interaction effects and standard errors in logit and probit models. *The Stata Journal*, 4(2):154–167.

Pagan, J. A. and Davila, A. (1997). Obesity, occupational attainment, and earnings. *Social Science Quarterly*, 78:757–770.

Poston, W., Ericsson, M., Linder, J., Haddock, C., Hanis, C., Nilsson, T., Astrom, M., and Foreyt, J. (1998). D4 dopamine receptor gene exon iii polymorphism and obesity risk. *Eating and Weight Disorders*, 3:71–77.

Quesenberry, C. P., Caan, B., and Jacobson, A. (1998). Obesity, health services use, and health care costs among members of a health maintenance organization. *Archives of Internal Medicine*, 158:466–472.

Register, C. A. and Williams, D. R. (1990). Wage effects of obesity among young workers. *Social Science Quarterly*, 71(1):130–141.

Robinson, P. (1988). Root-N-consistent semiparametric regression. *Econometrica*, 56:931–954.

Rosenzweig, M. R. and Shultz, T. P. (1983). Estimating a household production function: Heterogeneity, the demand for health inputs, and their effects on birth weight. *Journal of Political Economy*, 91:723–746.

Rubin, D. B. (1987). *Multiple Imputation for Nonresponse in Surveys.* Wiley.

Ruhm, C. J. (2007). Current and future prevalence of obesity and severe obesity in the united states. *Forum for Health Economics and Policy*, 10(2):1–26.

Sargent, J. D. and Blanchflower, D. G. (1994). Obesity and stature in adolescence and earnings in young adulthood. *Archives of Pediatric and Adolescent Medicine*, 148:681–687.

Schafer, J. (2000). *Analysis of Incomplete Multivariate Data.* Chapman & Hall/CRC.

Segal, N. and Allison, D. (2002). Twins and virtual twins: Bases of relative body weight revisited. *International Journal of Obesity*, 26:437–441.

Shimokawa, S. (2008). The labor market impact of body weight in China: a semiparametric analysis. *Applied Economics*, 40:949–968.

Staiger, D. and Stock, J. H. (1997). Instrumental variables regression with weak instruments. *Econometrica*, 65:557–586.

Stearns, P. N. (1997). *Fat History: Bodies and Beauties in the Modern West.* New York University Press.

Sunohara, G., Roberts, W., Malone, M., Schachar, R. J., and Tannock, R. (2000). Linkage of the dopamine d4 receptor gene and attention-deficit/hyperactivity disorder. *American Journal of Child and Adolescent Psychiatry*, 39:1537–1542.

Swanson, J., Oosterlaan, J., Murias, M., Schuck, S., Floodman, P., Spence, M. A., Wasdell, M., Ding, Y., Chi, H.-C., Smith, M., Mann, M., Carlson, C., Kennedy, J. L., Sargeant, J. A., Leung, P., Zhang, Y.-P., Sadeh, A., Chen, C., Whalen, C. K., Babb, K. A., Moyzis, R., and Posner, M. I. (2000). Attention deficit/hyperactivity disorder children with 7-repeat allele of the dopamine receptor d4 gene have extreme behavior but normal performance on critical neuropsychological test of attention. *Proceedings of the National Academy of Science*, 97:4754–4759.

Torgerson, J. S., Boldrin, M. N., Hauptman, J., and Sjöström, L. (2004). Xenical in the prevention of diabetes in obese subjects study. *Diabetes Care*, 27(1):155–161.

Van Buuren, S., Boshuizen, H., and Knook, D. (1999). Multiple imputation of missing blood pressure covariates in survival analysis. *Statistics in Medicine*, 18:681–694.

Wada, R. and Tekin, E. (2007). Body composition and wages. Working Paper 13595, National Bureau of Economic Research.

Wasserman, L. (2006). *All of Nonparametric Statistics.* Springer.

Wu, S. (2001). Adapting to heart conditions: the hedonic treadmill. *Journal of Health Economics*, 20(4):495–508.

Yang, Z., Gilleskie, D. B., and Norton, E. C. (2005). Prescription drugs, medical care, and health outcomes: A model of elderly health dynamics. Technical report, NBER Working Paper 10964.

Yatchew, A. (1998). Non-parametric techniques in economics. *Journal of Economic Literature*, 36:669–721.

Yatchew, A. (2003). *Semiparametric Regression for the Applied Econometrician*. Cambridge University Press.